

When and why defaults influence decisions: a meta-analysis of default effects

JON M. JACHIMOWICZ*

Columbia Business School, New York, NY, USA

SHANNON DUNCAN

Columbia Business School, New York, NY, USA

ELKE U. WEBER

Princeton University, Princeton, NJ, USA

ERIC J. JOHNSON

Columbia Business School, New York, NY, USA

Abstract: When people make decisions with a pre-selected choice option – a ‘default’ – they are more likely to select that option. Because defaults are easy to implement, they constitute one of the most widely employed tools in the choice architecture toolbox. However, to decide when defaults should be used instead of other choice architecture tools, policy-makers must know how effective defaults are and when and why their effectiveness varies. To answer these questions, we conduct a literature search and meta-analysis of the 58 default studies (pooled $n = 73,675$) that fit our criteria. While our analysis reveals a considerable influence of defaults ($d = 0.68$, 95% confidence interval = 0.53–0.83), we also discover substantial variation: the majority of default studies find positive effects, but several do not find a significant effect, and two even demonstrate negative effects. To explain this variability, we draw on existing theoretical frameworks to examine the drivers of disparity in effectiveness. Our analysis reveals two factors that partially account for the variability in defaults’ effectiveness. First, we find that defaults in consumer domains are more effective and in environmental domains are less effective. Second, we find that defaults are more effective when they operate through endorsement (defaults that are seen as conveying what the choice architect thinks the decision-maker should do) or endowment (defaults that are seen as reflecting the status quo). We end with a discussion of possible directions for a future research program on defaults,

* Correspondence to: Columbia Business School – Management Department, 3022 Broadway Uris Hall, Office 7-I, New York, NY 10027, USA. Email: jon.jachimowicz@columbia.edu

including potential additional moderators, and implications for policy-makers interested in the implementation and evaluation of defaults.

Submitted 21 August 2018; revised 4 December 2018;
accepted 10 December 2018

Introduction

When teaching students and practitioners about defaults – pre-selecting one choice option to increase the likelihood of its uptake – a figure that depicts the effect of defaults on organ donation often features prominently (Johnson & Goldstein, 2003). Organ donation defaults can be very simple, even consisting of the one-word difference between “If you want to be an organ donor, please check here” (opt-in) and “If you *don’t* want to be an organ donor, please check here” (opt-out). However, the ensuing difference in organ donation signup is dramatic, with percentages in the high nineties for opt-out countries and in the tens for opt-in countries (Johnson & Goldstein, 2003). These results seem to have influenced countries to change defaults: Argentina became an opt-out country in 2005 (La Nacion, 2005), Uruguay in 2012 (Trujillo, 2013), Chile in 2013 (Zúñiga-Fajuri, 2015) and Wales in 2015 (Griffiths, 2013), and The Netherlands and France will become opt-out countries by 2020 (Willsher, 2017; Leung, 2018).

The attractiveness of defaults as a choice architecture tool stems from their apparent effectiveness in a variety of different contexts and their relative ease of implementation. As a result, policy-makers and organizations regard defaults as a viable tool to guide individuals’ behaviors (Kahneman, 2011; Johnson *et al.*, 2012; Beshears *et al.*, 2015; Steffel *et al.*, 2016; Benartzi *et al.*, 2017). For example, one study showed that employees are 50% more likely to participate in a retirement savings program when enrollment is the default (i.e., they are automatically enrolled, with the option to reverse that decision) than when not enrolling is the default (Madrian & Shea, 2001). In response, Sen. Daniel Akaka (D-HI) introduced the Save More Tomorrow Act of 2012, which now provides opt-out enrollment in retirement savings for federal employees (Thaler & Benartzi, 2004; Akaka, 2012). Across many other domains and governments, defaults have also attracted increasing attention from policy-makers (Felsen *et al.*, 2013; Sunstein, 2015; Tannenbaum *et al.*, 2017).

However, the rise of defaults’ popularity should be seen in context: they are only a single tool in the choice architect’s toolbox (Johnson *et al.*, 2012). For example, while citizens could be defaulted into health insurance plans, they could also be asked to select their health insurance plan from a smaller,

curated choice set (Johnson *et al.*, 2013). Similarly, employees could be defaulted into retirement savings plans when joining a company, but alternatively, they could be given a limited time window in which to sign up (O'Donoghue & Rabin, 1999). Likewise, although consumers could be defaulted into more environmentally friendly automobile choices, gas mileage information could instead be presented in a more intuitive way to sway decisions toward more environmentally friendly options (Larrick & Soll, 2008; Camilleri & Larrick, 2014). Finally, instead of shifting toward an opt-out doctrine, policy-makers could also design active choice settings where individuals are required to make a choice (Keller *et al.*, 2011). Policy-makers thus have a large array of options to choose from, beyond defaults, when determining how to use choice architecture to attain desired outcomes.

Making an informed decision when selecting a choice architecture tool therefore requires information on how effective a tool is, as well as information about why a tool's effectiveness might vary. A maintenance worker's toolbox serves as a helpful analogy: to fix a problem, he or she must understand when the use of what tool may be more or less appropriate and how the tool should be handled to best address the underlying issue. However, because choice architects commonly do not have access to this information and are often inaccurate in their estimations of the default effect (Zlatev *et al.*, 2017), they may frequently fail to choose the most appropriate choice architecture tool or may deploy it inappropriately. In addition, in some cases, the implementation of an opt-out default may even reduce the take-up of the pre-selected option (Krijnen *et al.*, 2017). In fact, choice architects currently do not know how effective they can expect an implementation of a default to be, nor which factors in the design decisions may systematically alter how influential the application of a default may be. The current research seeks to address these issues by investigating how effective defaults are and when and why defaults' effectiveness varies.

We subsequently proceed as follows: we first present a meta-analysis of default studies that estimates the size of the default effect and its variability in prior research. We find that defaults have a sizeable and robust effect but that their effectiveness varies substantially across studies. We also investigate possible publication bias and find that – if anything – larger effect sizes are underreported.

We then explore the factors that may explain the observed variability of default effects in two different ways. We highlight that choice architects often make inadvertent decisions in studying default effects because they do not have perfect insight into which factors drive a default's effectiveness (Zlatev *et al.*, 2017). As a result, the variability in their design decisions allows us to investigate whether study factors systematically influence a

default's effectiveness, with the hope that our findings can subsequently inform the future and more deliberate design of defaults.

First, we examine whether study characteristics such as the choice domain or response mode can explain some of this variability, and we find that defaults that involve consumer decisions are more likely to be effective and defaults that involve environmental decisions are less likely to be effective. Second, we draw on an existing theoretical framework of default effects (Dinner *et al.*, 2011) to explore whether the variability in default effects could also be caused by differences in the mechanisms that may underlie the default effect in each study. Past research has demonstrated that defaults are multiply determined, depending on the extent to which they activate endorsement, endowment, and ease, and we find that both the nature and the number of mechanisms that are activated through the design of the default influence its effectiveness. We end with a discussion of possible directions for a future research program on defaults, including potential additional moderators, and implications for policy-makers interested in the implementation and evaluation of defaults.

Estimating the size and modeling the variability of default effects

We first aim to provide an estimate of the size and variability of default effects by conducting a meta-analysis of existing default studies. A meta-analysis combines the results of multiple studies to improve the estimate of an effect size by increasing the statistical power (Griffeth *et al.*, 2000; Judge *et al.*, 2002; Hagger *et al.*, 2010).

Inclusion criteria

We define the default effect as the difference in choice between the opt-out condition versus that in the opt-in condition. We include studies with both binary measures of choice (i.e., the percentage who choose the desired outcome in each condition) and continuous measures of choice (e.g., the average amount donated or invested in each condition). If a study has multiple relevant dependent measures, we include each measure as a separate observation. This is true of one study in our data, which looked at both the percentage who chose the desired outcome and their willingness to pay (Pichert & Katsikopoulos, 2008). If a study included multiple groups that should or could not be combined, an effect size is calculated for each. This is true of two studies in our data, one of which had two different pricing programs in their field study (Fowlie *et al.*, 2017) and another which looked at parents with different HPV vaccination intentions (Reiter *et al.*, 2012).

Because we define the default effect using opt-in and opt-out conditions, we focus only on studies that explicitly compare these two conditions. We exclude

any studies that explore defaults but do not contain a comparison between opt-in and opt-out conditions (e.g., they compare opt-out and forced choice or opt-in and forced choice). If a study investigates more conditions than just opt-in and opt-out (e.g., also includes forced choice), we only look at the data for the two relevant conditions. If a study looks at independent variables other than our two default conditions, we include only the effect of defaults on choice. Additionally, we exclude any studies for which missing information (such as means or standard deviations) prevents Cohen's d from being calculated. Finally, we include studies regardless of their publication date.

Data collection

We searched the EBSCO, ProQuest, ScienceDirect, PubMed and SAGE Publications databases and conference abstracts (Behavioral Decision Research in Management; Behavioral Science and Policy; Society for Judgment and Decision-Making; and Subjective Probability, Utility and Decision-Making) using the following search terms: 'Defaults' or 'Default Effect' or 'Advance Directives' or 'Opt-out' or 'Opt-in' AND 'Decisions' or 'Decision-Making' or 'Environmental Decisions' or 'Health Decisions' or 'Consumer Behavior'. We also sent requests for papers to two academic mailing lists: SJDM and ACR. Our search concluded in May 2017.

In total, we found 58 datasets from 55 studies included in 35 articles that fit our inclusion criteria ($n = 73,675$, ranging from 51 to 41,952¹). These articles come from a variety of journals, including, but not limited to, *Science*, *Journal of Marketing Research*, *Journal of Consumer Psychology*, *Medical Decision Making*, *Journal of Environmental Psychology* and *Quarterly Journal of Economics*. The meta-analysis data and code are publicly available via the Open Science Framework (<https://osf.io/tcbh7>).

Effect size coding

To combine all individual studies into one meta-analytic model, we calculate the Cohen's d for the differences between the opt-out and the opt-in conditions. We code effects so that a positive d -value is associated with greater choice in the opt-out condition and a negative d -value is associated with greater choice in the opt-in condition. For dependent variables that are measured on a continuous scale (e.g., the amount of money donated or invested), we calculate Cohen's d as the difference between the means, divided by the pooled standard

¹ The study with 41,952 observations was Ebeling and Lotz (2015), a randomized controlled trial of German households with a nationwide energy supplier.

deviation (Cohen, 1988). For dependent variables that are measured on a binary scale, we calculate the Cohen's d using an arcsine transformation (Lipsey & Wilson, 2001; Scheibehenne *et al.*, 2010; Chernev *et al.*, 2012).

Due to the high level of variation in the results of our selected studies, we use a random-effects model via the restricted maximum likelihood estimator method. All analyses were conducted in R version 1.1.383 using the 'metafor' package (Viechtbauer, 2010). The studies are weighted using inverse-variance weights, which has been shown to perform better than weighting by sample size in random-effects analysis (Marín-Martínez & Sánchez-Meca, 2010).

Since the data are nested (58 observations from 55 studies in 35 articles), we also use a random-effects model that accounts for three levels: individual observations; observations within the same study (either separate groups from the same study or multiple dependent measures from the same study); and studies within the same article. By using these three levels, we can take into account that observations derived from the same study or article are likely to be more similar than observations from different studies or articles (Rosenthal, 1995; Thompson & Higgins, 2002; Sánchez-Meca *et al.*, 2003; Chernev *et al.*, 2012).

Results: effect size

Our analysis reveals that opt-out defaults lead to greater uptake of the pre-selected decision than opt-in defaults ($d = 0.68$, 95% confidence interval [CI] = 0.53–0.83], $p < 0.001$), producing a medium-sized effect given conventional criteria (Cohen, 1988). This is robust to running the model that accounts for the three levels: the observation level, the study level and the article level ($d^+ = 0.63$, 95% CI = 0.47–0.80], $p < 0.001$). In other words, when we account for observations within the same study, those that come from separate groups or different dependent measures, as well as studies that come from the same articles, our results largely do not differ. In comparison to a decision where participants must explicitly give their consent to follow through with a desired course of action, a decision with a pre-selected option increases the likelihood that the option is chosen by 0.63–0.68 standard deviations. [Figure 1](#) illustrates this result in a forest plot.

Binary studies

We also examine the Cramér's V for all binary dependent measure observations in our analysis – a measure of association for nominal values that gives a value between 0 (no association between variables) and 1 (perfect association between variables) – which we calculate by taking the square root of the

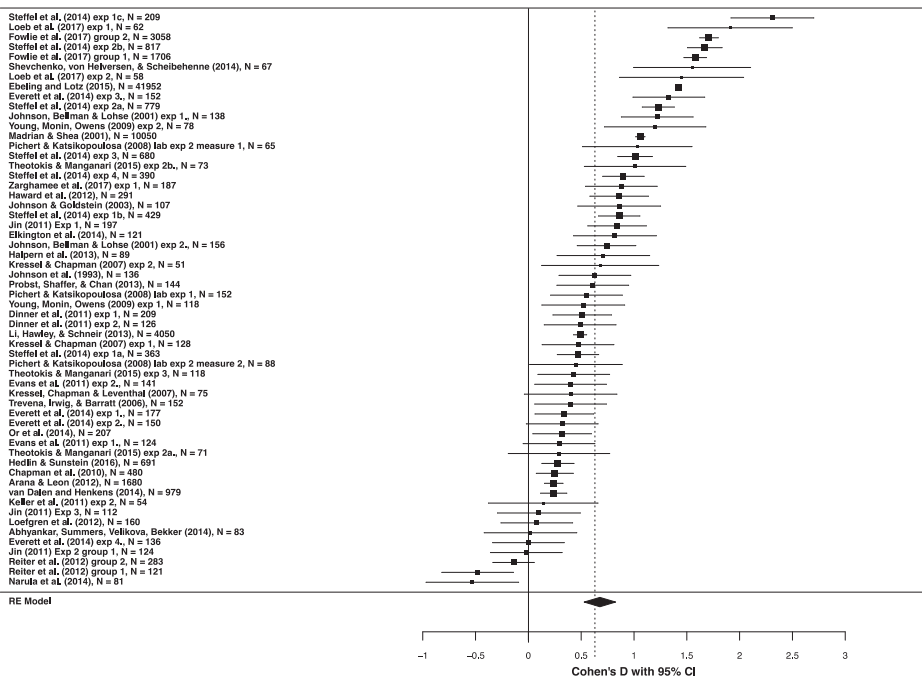


Figure 1. Forest plot of default effect size (all studies)

Notes: Each line represents one observation. The position of the square depicts the effect size; the size of the square, the weighted variance; and the line through each square, the confidence interval (CI) for each observation. The vertical dotted line represents the weighted averaged effect size
 RE = random effects

chi-squared statistic divided by the sample size and the minimum dimension, minus 1 (Cramér, 1946). We note that this calculation is not new or different information, but merely a translation of the Cohen’s *d* results to a different scale for interpretation purposes for binary choice datasets (38 out of 58). We again find that opt-out defaults lead to significantly greater uptake of the pre-selected decision than opt-in defaults ($V^+ = 0.29$, 95% CI = 0.21–0.37], $p < 0.001$) by an absolute average of 27.24%. Hence, our meta-analysis indicates that defaults, in aggregate, have a considerable influence on decision-making outcomes.

Publication bias

We next estimate the extent of publication bias in the published defaults literature (see also Duval & Tweedie, 2000; Carter & McCullough, 2014; Franco

et al., 2014; Simonsohn *et al.*, 2014; Dang, 2016). It is possible that there is a file-drawer problem, in which non-significant default studies are not published. To investigate this, we first create a funnel plot that plots the treatment effect of studies in a meta-analysis against a measure of study precision: in this case, Cohen's d as a function of the standard error (see Figure 2). Each black dot in Figure 2 represents an effect size. Higher-powered studies are located higher, and lower-powered studies are located lower. In the absence of publication bias, studies should be distributed symmetrically, depicted by the white shading in Figure 2. Reviewing the funnel plots highlights that several observations appear outside of the funnel on both sides, suggesting potential publication bias (Duval & Tweedie, 2000).

We next conduct the trim-and-fill method, an iterative nonparametric test that attempts to estimate which studies are likely missing for a variety of reasons, such as publication bias, but also including other forms of bias (such as poor study design; Duval & Tweedie, 2000). In simple terms, this method investigates which effect size estimates are missing, since, in the absence of any form of bias, the funnel should be symmetric. This analysis reveals that eight studies are missing from the funnel plot (represented by the white dots in Figure 3). Including these studies *increases* the overall effect to $d^* = 0.80$, 95% CI = 0.65–0.96, $p < 0.001$; this estimate remains directionally the same as our prior analysis and is significantly different from zero. This indicates that, if anything, default studies finding larger effects are missing from the literature. However, because Egger *et al.*'s (1997) test for asymmetry of the funnel plot is not significant ($t(56) = -0.39$, $p = 0.69$), the likely absence of studies does not lead to inadequate estimation of the default effect. While this result is encouraging, Egger's regression is prone to Type I errors in cases where heterogeneity is high (Sterne *et al.*, 2011), as is the case in the current meta-analysis. These results should thus be interpreted with caution.

Why do default effects vary?

While Figure 1 shows a sizeable average default effect, it also highlights significant variation in the effect size. Even by visually assessing the effect sizes in Figure 1, it becomes apparent that the default effect size varies widely: 46 observations find a statistically significant and positive effect (i.e., the observations are to the right of 0 and the confidence interval excludes 0), ten observations do not find a statistically significant effect (i.e., the confidence interval includes 0) and two observations find a statistically significant and negative effect (i.e., the observations are to the left of 0 and the confidence interval excludes 0).

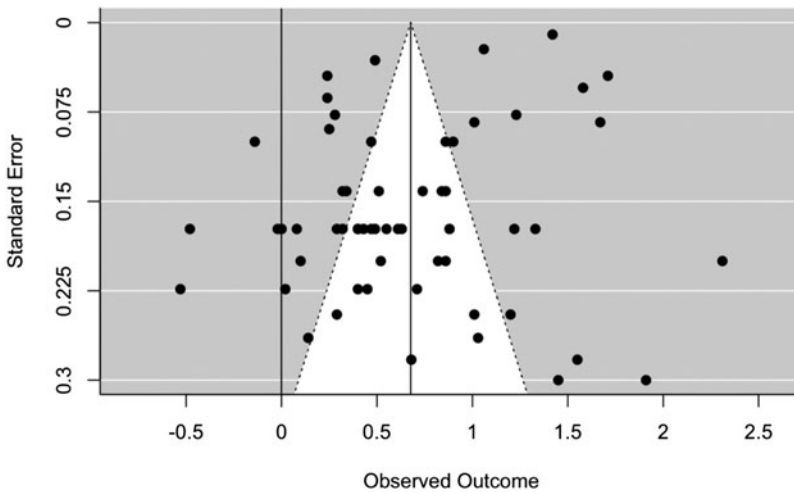


Figure 2. Funnel plot of individual effect sizes

Notes: Each black dot represents an effect size. Higher-powered studies are located higher, and lower-powered studies are located lower. The x -axis depicts the effect size, with the black line in the middle representing the average effect size. The plot should ideally resemble a pyramid (shaded white), with scatter that arises as a result of sampling variation

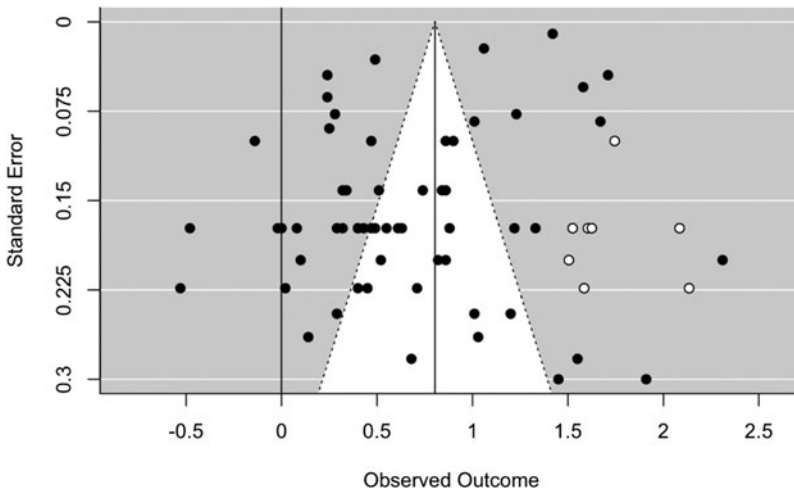


Figure 3. Trim-and-fill funnel plot

Notes: Each black dot represents a study. The white dots represent missing studies. The black line in the middle represents the average effect size

To quantify the extent of the variability of the default effect, we conduct analyses that assess this heterogeneity using the I^2 statistic, a measure that reflects both the variability in the default effect and the variability in the sampling error. In our base model, we find an I^2 of 98.21%. We employ methods that apply the use of I^2 to multilevel meta-analytic models (Nakagawa & Santos, 2012) and find an I^2 of 98.01% for our three-level model (observation level, study level, and article level), which is considered to be very high heterogeneity (Higgins *et al.*, 2003). This result is consistent with other analyses that find that the heterogeneity of effect sizes tends to increase as the effect size increases (Klein *et al.*, 2014).

We further refine this analysis to distinguish between-cluster heterogeneity from within-cluster heterogeneity (Cheung & Chan, 2014). We do this because our model contains multiple variance components: for the article level (between-cluster) and for observations within the same studies and studies within the same articles (within-cluster heterogeneity). Parceling out these distinct sources, we find that 30.21% of the heterogeneity is at the article level, 63.58% is at the studies within-articles level, 4.21% is at the observations within-studies level, and the remaining 2.00% is due to sampling variance. This analysis suggests that there is significant variability in the size of default effects.

Given this variation in the size of defaults effect, we next explore potential explanations for it. We specifically examine two potential factors: (1) Do defaults differ because of the characteristics of the studies? (2) Do default studies that use different mechanisms produce different-sized default effects?

Do characteristics of the studies explain the default effect size?

To investigate whether characteristics of default studies partially explain some of the differences in effect sizes, we use methods from prior meta-analyses to assess additional study attributes (e.g., Carter *et al.*, 2015).

Study characteristics

Domain

We first code each study into three main types of domain: ‘environmental’ (‘0’ for non-environmental and ‘1’ for environmental, defined as making a choice that is related to pro-environmental behavior), ‘consumer choice’ (‘0’ for non-consumer choice and ‘1’ for consumer choice, defined as decisions related to buying a product or service) and ‘health’ (‘0’ for non-health and ‘1’ for health, defined as making a choice related to health care treatment, organ donation or health behaviors). We note that ten studies are coded as being in

more than one domain (e.g., consumer choice of environmental products). The first two authors of the current manuscript coded each study's domain, and interrater reliability was high (Cohen's $\kappa = 0.94$).

Field experiment

We next code for the type of study it was: a field experiment, coded as '1', or lab experiment, coded as '0'. We did so in order to determine whether studies in a real-world choice setting varied in default effect size in comparison to those in a hypothetical choice setting (online or in person). Some have suggested that lab experiments should have larger effect sizes due to a larger amount of control over the experiment (Cooper, 1981). However, others have found that field studies can elicit larger effect sizes than lab studies, in part because they are often preceded by a viability study, making those field studies that are conducted more likely to find a stronger effect (Peterson *et al.*, 1985).

Location

We then code for the study location (i.e., whether it was conducted in the USA, coded as '1', or not in the USA, coded as '0'). In other words, this coding was conducted to explore whether the location of the participants who took part in the study influenced the default effect (see also Cadario & Chandon, 2018).

Time of publication

We also code for the decade in which a paper was published, with '1' being the 1990s, '2' being the 2000s and '3' being the 2010s. We specifically code for decade to determine whether default effect sizes have changed over the time that they have been studied, as effect sizes in published research frequently decrease over time (Szucs *et al.*, 2015).

Response mode

We characterize the dependent variables as binary ('yes' or 'no' choice), coded as '1', or continuous (e.g., the amount of money invested or donated), coded as '0'. Given that these dependent variables involve a different type of choice, we aim to determine whether the default effect size varies based on which type of choice is made.

Sample size

We also code each observation for the total sample size, with '0' reflecting a sample size below 1000 and '1' reflecting a sample size above 1000. While effect size calculations should be independent of sample size, we explore whether the size of the default effect varies with the sample size across the studies. Since studies are more likely to be published if they find a statistically

significant effect and it takes a larger effect size to achieve statistical significance in a small study than a large one, small studies may be more likely to be published if their effects are large (Slavin & Smith, 2009). Additionally, small studies tend to be of lesser methodological quality, leading to greater variability from studies with smaller sample sizes, which could introduce a higher probability of positive effect sizes (Kjaergard *et al.*, 2001).

Presentation mode

We also distinguish between studies where the default is presented online, coded as ‘1’, or not presented online, coded as ‘0’. For example, some defaults are presented via an in-person form, such as a default for a carbon-emission offset on a form, whereas others are presented via the internet, such as a default for a product selection while shopping online. We code for this to examine whether differences in presentation mode influence default efficacy.

Benefits self vs. others

To explore whether differences in who benefits from the choice influences the size of the default effect, we code for differences in the nature of the choice facing the participants; that is, whether the choice would be more beneficial to the self or to others. Choices that benefited the self more were coded as ‘1’ and choices that benefited others more were coded as ‘0’.

Financial consequence

Finally, to explore whether choices that involved a financial consequence would alter the default effect, we code for whether the choice that the participant made resulted in an actual financial consequence (i.e., donating a portion of their participation reimbursement to charity). Studies that included a financial consequence of choice were coded as ‘1’ and those that did not were coded as ‘0’.

Results: study characteristics

We add study characteristics as moderators to the prior random-effects model. For this model, only the regression coefficient for consumer domains ($b = 0.73$, $SE = 0.23$, $p = 0.003$; see Table 1) is statistically significant and positive, while the regression coefficient for environmental domains is marginally significant and negative ($b = -0.47$, $SE = 0.27$, $p = 0.08$). Including study characteristics as moderators reduces the heterogeneity by 4.67% to $I^2 = 93.54\%$. Given that we extracted multiple effect sizes from some of the studies, we also reran the analysis using robust variance estimation using the ‘clubSandwich’

Table 1. Model results including study characteristics

Effect	<i>b</i>	SE	<i>t</i>	p	Robust SE	p
Intercept	0.82	0.45	1.82	0.08	0.56	0.18
<i>Study characteristics</i>						
Binary response	−0.10	0.15	−0.65	0.52	0.18	0.59
Decade of study	−0.13	0.16	−0.82	0.41	0.20	0.53
Sample size	0.29	0.26	1.14	0.26	0.22	0.22
Online vs. offline	−0.05	0.16	−0.32	0.74	0.15	0.72
Location: USA vs. other	−0.08	0.23	−0.36	0.72	0.26	0.76
Benefits self	0.14	0.17	0.82	0.42	0.17	0.43
Field experiment	−0.02	0.18	−0.09	0.93	0.20	0.94
Financial consequence	0.34	0.21	1.62	0.11	0.25	0.19
<i>Domain</i>						
Environmental	−0.47	0.27	−1.76	0.08	0.26	0.09
Consumer	0.73	0.23	3.15	0.003	0.28	0.03
Health	0.001	0.20	0.006	0.99	0.25	0.99
Observations	58					
Studies	55					
R ²	0.24					

package (Pustejovsky, 2015). However, the results of the analyses did not meaningfully change when using this estimation.

Do study characteristics explain the variation in default effect size? Our analysis suggests that defaults are more effective in consumer domains and that the default effect may be weaker in environmental domains. No other study characteristic explained any further systematic variance in the variability of default effects present in prior studies. We next investigate whether the presence or absence of different mechanisms known to produce default effects partially explains why a default's effectiveness varies across studies.

Do different channels explain variation in default effect size?

A theory-based approach to meta-analyses suggests that insights from prior research can inform which factors may account for variation and can indicate how to investigate if the observed outcomes are following an expected pattern (Becker, 2001; Higgins & Thompson, 2002). We use prior research to further investigate factors that explain the variation in defaults' effectiveness. In particular, we draw on the framework developed by Dinner *et al.* (2011), who propose that defaults influence decisions through three psychological channels – endorsement, ease, and endowment – that can play a role both individually and in parallel (drawing on prior research, e.g., McKenzie *et al.*, 2006, on

endorsement; and Choi *et al.*, 2002, on the path of least resistance). That is, decision-makers are more likely to choose the pre-selected option because: (1) they believe that the intentions of the choice architect, as suggested through the choice design, are beneficial to them; (2) they can exert less effort when staying with the pre-selected option; and/or (3) they will evaluate other options in reference to the pre-selected option with which they are already endowed (Dinner *et al.*, 2011).

One interpretation of these findings is that defaults are more effective when they activate more channels; that is, the effects of the underlying mechanisms are additive. Similarly, defaults may be less effective – or may not influence decisions at all – if they activate fewer of the psychological channels in the minds of decision-makers. Because choice architects may not have perfect knowledge of the underlying drivers of defaults' effectiveness, it is likely that there are systematic differences in the design of defaults and the activation of the channels driving its effects (Zlatev *et al.*, 2017). We intend to exploit this occurrence and use it to evaluate the relative importance of each underlying driver to the default effect. We next describe each channel in more detail and describe how we code each study for the strength of each mechanism. Our aim is to examine whether the variation in default effects is partially driven by the extent to which different defaults activate these three channels.

The three channels: endorsement, ease, and endowment

Individuals commonly perceive defaults as conveying an endorsement by the choice architect (McKenzie *et al.*, 2006). As a result, a default's effectiveness is in part determined by whom decision-makers perceive to be the architect of the choice and by what their attitudes toward this perceived choice architect are. For example, one study finds that defaults are less effective when individuals do not trust the choice architect because the individuals believe that the choice design was based on intentions differing from their own (Tannenbaum *et al.*, 2017). Endorsement is thus one mechanism that drives a default's effectiveness: the more decision-makers believe that the default reflects a trusted recommendation, the more effective the default is likely to be.

Decision-makers may also favor the defaulted choice option because it is easier to stay with the pre-selected option than to choose a different option. When an option is pre-selected, individuals may not evaluate every presented option separately, but rather may simply assess whether the default option satisfies them (Johnson *et al.*, 2012). In addition, different default designs differ in how easy it is for the decision-maker to change away from the default; when more effort is necessary to switch away from the pre-selected option, decision-makers may be more likely to stick with the default. Ease is

thus a second mechanism that drives a default's effectiveness: the harder it is for decision-makers to switch away from the pre-selected option, the more effective the default is likely to be.

A third channel that drives the effectiveness of a default is endowment, or the extent to which decision-makers believe that the pre-selected option reflects the status quo. The more decision-makers feel endowed with the pre-selected option, the more likely they are to stay with the default as a result of reference-dependent encoding and loss aversion (Kahneman & Tversky, 1979). For example, one study finds that arbitrarily labeling a policy option as the 'status quo' increases the attractiveness of that option (Moshinsky & Bar-Hillel, 2010). Endowment is thus a third mechanism that drives a default's effectiveness: the more decision-makers feel that the default reflects the status quo, the more effective the default is likely to be.

Coding the activation of the three default channels

It is not straightforward to identify whether the variation in default effects is explained by the activation of each of the three default channels. Ideally, we would have access to study respondents' ratings of the choice architect to evaluate endorsement (as collected by Tannenbaum *et al.*, 2017, and Bang *et al.*, 2018), measures of reaction time to evaluate ease (as collected by Dinner *et al.*, 2011) and measures of thoughts to evaluate endowment (as collected by Dinner *et al.*, 2011). However, these data are not available in the vast majority of default studies.

In the absence of such information, we trained two coders – a graduate student and a senior research assistant, who are not part of the author team – to rate each default study on the extent to which its design likely triggered each of the three channels (endorsement, ease and endowment; see Cadario & Chandon, 2018, and Jachimowicz, Wihler *et al.*, 2018, for similar approaches). The two coders were first trained with a set of default studies that did not meet the inclusion criteria and next coded each default study on each of the three channels. Endorsement and ease were coded on a scale ranging from '0' (this channel did not play a role) to '1' (this channel played somewhat of a role) and '2' (this channel played a role), rated on half-steps. For endowment, in trialing the coding scheme, we recognized that a scale of '0' (this channel did not play a role) and '1' (this channel played a role) is more appropriate. Appendix A contains a detailed description of the coding scheme. Interrater reliability, calculated via Cohen's κ , was acceptable for endorsement ($\kappa = 0.59$), ease ($\kappa = 0.58$) and endowment ($\kappa = 0.80$; Landis & Koch, 1977). Correlations between channels are not statistically significant (see Appendix B for scatter plots and correlations).

Results: default channels

We subsequently add the coding for default channels to the prior random-effects model as moderators to evaluate whether these partially account for the variability of default effects. The analysis reveals that, as predicted, both endorsement ($b = 0.32$, $SE = 0.15$, $p = 0.038$) and endowment ($b = 0.31$, $SE = 0.15$, $p = 0.044$) are significant moderators of the default effect (see Table 2). Contrary to our prediction, ease is not a significant moderator ($b = -0.05$, $SE = 0.15$, $p = 0.75$). The addition of the coding for default channels further reduces heterogeneity to $I^2 = 92.32\%$.

As in the previous model, consumer domains remain statistically significant and positive ($b = 0.89$, $SE = 0.23$, $p = 0.0003$), and the environmental domain is now statistically significant and negative ($b = -0.60$, $SE = 0.26$, $p = 0.028$). We also re-ran the model using robust variance estimation using the ‘clubSandwich’ package (Pustejovsky, 2015), and we find that in this analysis the endowment channel ($b = 0.31$, $SE = 0.16$, $p = 0.081$) and the environmental domain ($b = -0.60$, $SE = 0.28$, $p = 0.056$) drop to marginal significance; all other results hold in this specification. Finally, we test for multicollinearity by examining the correlation matrix of the independent variables and do not find evidence for multicollinearity, as all correlations were small to moderate, and most were nonsignificant.²

General discussion

Defaults have become an increasingly popular policy intervention, and rightly so, given that our meta-analysis shows that defaults exert a considerable influence on individuals’ decisions: on average, pre-selecting an option increases the likelihood that the default option is chosen by 0.63–0.68 standard deviations, or a change of 27.24% in studies that report binary outcomes. If anything, our publication bias analyses highlight that larger effect sizes are underreported, suggesting that researchers may not bother to report replications of what are believed to be strong effects. While it is difficult to compare the effectiveness of defaults with other interventions outside of one focal study, we note that the effect for defaults we find in the current meta-

² In contrast to many other meta-analytic studies, our set of studies contain very large field studies, containing tens of thousands of observations. As a robustness check, we also conduct additional heterogeneity analyses excluding observations where the sample size was above 1000. In the base model (without moderators), we find an I^2 of 92.85%; when adding study characteristics, we find an I^2 of 90.47%; and when additionally entering the default channel coding, we find an I^2 of 89.12%. These additional analyses thus highlight that the extent of the heterogeneity is at least partially driven by the large sample sizes included in the meta-analysis.

Table 2. Model results including default channels

Effect	<i>b</i>	SE	<i>t</i>	<i>p</i>	Robust SE	<i>p</i>
Intercept	0.36	0.49	0.74	0.46	0.62	0.57
<i>Default channels</i>						
Endorsement	0.32	0.15	2.14	0.04	0.13	0.02
Ease	-0.05	0.15	-0.32	0.75	0.14	0.74
Endowment	0.31	0.15	2.07	0.04	0.16	0.08
<i>Study characteristics</i>						
Binary response	-0.18	0.15	-1.17	0.25	0.17	0.30
Decade of study	-0.16	0.16	-1.03	0.31	0.22	0.48
Sample size	0.36	0.25	1.48	0.15	0.24	0.17
Online vs. offline	0.02	0.16	0.09	0.93	0.13	0.91
Location: USA vs. other	-0.12	0.22	-0.58	0.57	0.26	0.64
Benefits self	0.22	0.17	1.28	0.21	0.16	0.18
Field experiment	0.05	0.17	0.28	0.78	0.17	0.77
Financial consequence	0.39	0.21	1.85	0.07	0.29	0.21
<i>Domain</i>						
Environmental	-0.60	0.26	-2.27	0.03	0.28	0.06
Consumer	0.89	0.23	3.91	0.00	0.29	0.01
Health	0.25	0.23	1.14	0.26	0.26	0.34
Observations	58					
Studies	55					
R ²	0.31					

analysis is considerably larger than a recent meta-analysis evaluation of healthy eating nudges (including plate size changes or nutrition labels, $d=0.23$; Cadario & Chandon, 2018), a recent meta-analysis of the effect of Opower's descriptive social norm intervention on energy savings ($d=0.32$; Jachimowicz, Hauser *et al.*, 2018), as well as a meta-analysis of framing effects on risky choice ($d=0.31$; Kühberger, 1998). Thus, defaults constitute a powerful intervention that can meaningfully alter individuals' decisions.

In addition, our analysis also reveals that there is substantial variation in the effectiveness of defaults, which indicates that a choice architect who deploys a default may have difficulty estimating the effect size he or she can expect from an implementation of defaults; this effect size may be substantially lower or higher than the meta-analytic average. This complicates the implementation of defaults by policy-makers, who, in order to decide which choice architecture tool to use, would like to know how large a default effect they can expect (Johnson *et al.*, 2012; Benartzi *et al.*, 2017). We note that this variation is likely driven by choice architects' imperfect understanding of the consequences of differences in default designs (Zlatev *et al.*, 2017).

To better understand the effectiveness of defaults and to enable policy-makers to better consider when and how to use defaults, we next examined factors that may at least partially explain the variation in defaults' effectiveness, drawing on an earlier framework proposed by Dinner *et al.* (2011) and empirically supported by several subsequent studies (e.g., Tannenbaum *et al.*, 2017; Bang *et al.*, 2018). Such an analysis is complicated by the fact that both study characteristics and potential mechanisms do not reflect systematic variation, but rather reflect the decisions of researchers about what studies to conduct and how to implement the default intervention. In addition, this analysis relies on our coders' ability to identify which channel is activated, which may be called into question. With this caveat, we believe that there are two substantial insights to be gained from our analysis.

First, there are domain effects worth exploring. We find that consumer domains show larger default effects and environmental domains have smaller default effects. We can only speculate about why this occurs, identifying it as a question that awaits further research. Perhaps consumer preferences are less strongly held than preferences in other domains and environmental preferences more strongly – a hypothesis described in more detail in the 'Limitations and future directions' section below. Second, we show that if the design of the default activates two of the three previously hypothesized channels of defaults' effectiveness (Dinner *et al.*, 2011), there is a significant increase in the size of the default effect.

However, we urge caution in interpreting these results, including the absence of an effect for ease. The ways that the three channels are measured in the current research are only noisy approximations, which attenuates our ability to detect systematic differences in the variability in defaults' effectiveness. While our coding provides a tentative examination of these issues, this approach is useful only because the underlying mechanisms have not been measured in the vast majority of prior default studies. In particular, ease does not seem to be systematically manipulated in studies, but is often varied in real-world applications. For example, Chile changed the means of opting out of being an organ donor from checking a box during the renewal of national identity cards to requiring a notarized statement (Zúñiga-Fajuri, 2015), thus increasing the difficulty of switching away from the default. The set of studies included in our meta-analysis lacks this kind of variation, which may partially account for the lack of a statistically significant effect for ease.

For policy-makers, our findings contain an important lesson: design choices that may have come about inadvertently and may seem inconsequential can have substantial consequences for the size of the default effect. As a result, choice architects may want to more systematically consider the extent to which, for example, the decision-maker believes that the choice architect has

their best interests in mind (endorsement), or to what extent decision-makers believe that the default is the status quo (endowment). Indeed, what may reflect more inadvertent decisions could turn into systematically made design decisions that could make the future implementation of defaults more successful. We next detail further necessary changes to help ensure this outcome is achieved.

Limitations and future directions

Defaults are often easy to implement, and this creates a temptation: to influence behavior, choice architects may prefer to set a default over other choice architecture tools. However, defaults vary in their effectiveness, and setting a default may not always be the most suitable intervention. In addition, choice architects are often inaccurate in their estimations of the default effect (Zlatev *et al.*, 2017). Our analysis underscores that when implementing defaults, one must test applications rigorously, rather than just assuming that they will always work as expected (Jachimowicz, 2017).

Our findings also suggest that future default studies should include measures of the default channels (endorsement, ease, and endowment). Where possible, choice architects could assess how decision-makers evaluate the choice architect's intentions, how easy decision-makers felt it was to opt out, or to what extent decision-makers believed that the default reflected the status quo. Ideally, these three mechanisms should be measured or manipulated systematically to better understand the size of default effects and the influence of context. For example, to test the endorsement channel of default effects, future research could systematically manipulate the source of who instituted the default (e.g., their status, purpose, etc.). In addition, we call on future studies to make more detailed information – including the original stimuli – publicly available in order to further help us to understand which channels may be driving a particular default's effectiveness. We note that studies have begun to explore the causal effects of these mechanisms, and we echo the call for future research to further advance this direction, especially as to how these effects may play out across different domains (e.g., Dinner *et al.*, 2011; Tannenbaum *et al.*, 2017; Bang *et al.*, 2018).

In addition, because choice architects have many ways of influencing choices beyond defaults (Johnson *et al.*, 2012), we call on future research to evaluate defaults relative to alternative choice architecture tools. In our introduction, we describe the analogy of a worker's toolbox, who must understand when the use of what tool may be more or less appropriate. While gaining a deeper appreciation of the effect size and reasons underlying the variability of default effects is a first important step toward this end, we note that

future research comparing default effects to alternative choice architecture interventions is a necessary complement. Such research would allow future studies to provide further insight into when and how defaults are more likely to exert a larger effect on decisions and in what cases policy-makers and other choice architects should rely on other tools in the toolbox (Johnson *et al.*, 2012; Benartzi *et al.*, 2017).

In addition, the effectiveness of defaults is particularly important given that prior research finds that public acceptance of choice-architecture interventions rests in part on their perceived effectiveness (Bang *et al.*, 2018; Davidai & Shafir, 2018). That is, an increase in the perceived effectiveness of choice-architecture interventions makes others view the intervention as more acceptable. To improve rates of acceptance of choice-architecture interventions more broadly, and of defaults more specifically, future studies could explore how the communication of the default effect found in the meta-analysis presented here would influence the evaluation of their further implementation. An application of defaults across policy-relevant domains may therefore rest on the communication of their effectiveness (Bang *et al.*, 2018; Davidai & Shafir, 2018).

We also propose additional variables that may moderate the default effect but could not be included in the current study due to a lack of available data, and we call on future research to either measure or manipulate these variables. One such variable is the intensity of a decision-maker's underlying preferences – what one might call 'preference strength'. When individuals care deeply about their inclination regarding a particular choice, they are more likely to have thought about their decisions and to be resistant to outside influence (Eagly & Chaiken, 1995; Crano & Prislin, 2006). In other words, defaults may be less likely to influence those who have strong preferences. Our finding that defaults in consumer domains are more effective and defaults in environmental domains less effective could in part be explained by this perspective, as preferences for consumption may be less strongly held than preferences for environmental choices.

A closely related but distinct moderator may focus on how important a particular decision is to an individual – what one might call 'decision importance'. That is, while individuals may believe that a particular decision is important, they may not have strongly formed preferences to help inform them how to respond. In cases where decision importance is high, individuals may be especially motivated to seek out novel information or otherwise exert effort to ascertain their decision. As a result, defaults that operate primarily through the ease channel may be less likely to have an effect in such circumstances, as individuals will be more motivated to exert effort.

Another important factor may be the distribution of underlying preferences. In some cases, the population of decision-makers may largely agree on what

they want; in other cases, they may vastly differ in opinion. This perspective is built into the design of defaults, which are based on the assumption that they allow those whose preferences are different from the default option to easily select an alternative (Thaler & Sunstein, 2008). However, this also suggests that defaults may be less effective in settings where preferences vary more widely than in instances where individuals' preferences diverge less, as the prevalence of decision-makers who disagree with the default is higher. We note that there may also be cases where the variance in underlying preferences is low, but because they are misaligned with the default, it may also be less effective.

Future research could therefore further investigate how the underlying preferences of the population presented with the default shape the default's effectiveness. That is, researchers and policy-makers interested in deploying defaults may have to consider what the distribution of decision-makers' underlying preferences is and how strongly these individuals hold such preferences. This could be done by including a forced-choice condition to assess what occurs in the absence of defaults, which would also allow the choice architect to see what the distribution of preferences might be in the absence of the intervention. We note that one consequence of a better understanding of the heterogeneity of underlying preferences could be the design of 'tailor-made' defaults, whereby the pre-selected choice differs as a function of the decision-makers' likely preferences (Johnson *et al.*, 2013). Evaluating the intended population's preferences may therefore reflect a crucial component in deciding when to deploy defaults.

Conclusion

On average, defaults exert a considerable influence on decisions. However, our meta-analysis also reveals substantial variability in defaults' effectiveness, suggesting that both *when* and *how* defaults are deployed matter. That is, both the context in which a default is used and whether the default's design triggers its underlying channels partially explain the variability in the default effect. To design better defaults in the future, policy-makers and other choice architects should consider this variability of default studies and the dynamics that may underlie it.

References

References marked with * are included in the meta-analysis.

- * Abhyankar, P., B. A. Summers, G. Velikova and H. L. Bekker (2014), 'Framing options as choice or opportunity: does the frame influence decisions?', *Medical Decision Making*, 34(5): 567–582.

- Akaka, D. K. (2012), Save More Tomorrow Act of 2012. Retrieved from http://www.federalnewsradio.com/wp-content/uploads/pdfs/Akaka_TSP_bill_statement.pdf
- * Araña, J. E. and C. J. León (2013), 'Can defaults save the climate? Evidence from a field experiment on carbon offsetting programs', *Environmental and Resource Economics*, 54(4): 613–626.
- Bang, H., S. Shu and E. U. Weber (2018), 'The role of perceived effectiveness on the acceptability of choice architecture', *Behavioural Public Policy*, 1–21.
- Becker, B. J. (2001), 'Examining theoretical models through research synthesis: The benefits of model-driven meta-analysis', *Evaluation & the Health Professions*, 24(2): 190–217.
- Benartzi, S., J. Beshears, K. L. Milkman, C. R. Sunstein, R. H. Thaler, M. Shankar, ... S. Galing (2017), 'Should Governments Invest More in Nudging?' *Psychological Science*, 28(8): 1041–1055.
- Beshears, J., J. J. Choi, D. Laibson, B. C. Madrian and S. Wang (2015), Who Is Easier to Nudge? NBER Working Paper, 401.
- Brown, C. L., and A. Krishna (2004), 'The Skeptical Shopper: A Metacognitive Account for the Effects of Default Options on Choice', *Journal of Consumer Research*, 31(3): 529–539.
- Cadario, R., and P. Chandon (2018), 'Which Healthy Eating Nudges Work Best? A Meta-Analysis of Field Experiments', *Marketing Science*.
- Camilleri, A. R., and R. P. Larrick (2014), 'Metric and Scale Design as Choice Architecture Tools', *Journal of Public Policy & Marketing*, 33(1): 108–125.
- Carter, E. C., and M. E. McCullough (2014), 'Publication bias and the limited strength model of self-control: has the evidence for ego depletion been overestimated?' *Frontiers in Psychology*.
- Carter, E. C., L. M. Kofler, D. E. Forster and M. E. McCullough (2015), 'A Series of Meta-Analytic Tests of the Depletion Effect: Self-Control does not Seem to Rely on a Limited Resource', *Journal of Experimental Psychology: General*, 144(4): 796–815.
- * Chapman, G. B., M. Li, H. Colby and H. Yoon (2010), 'Opting in vs opting out of influenza vaccination', *JAMA*, 304(1): 43–44.
- Chernev, A., U. Bockenholt and J. Goodman (2012), 'Choice overload: A conceptual review and meta-analysis', *Journal of Consumer Psychology*, 25(2): 333–358.
- Cheung, S. F., and D. K. S. Chan (2014), 'Meta-analyzing dependent correlations: An SPSS macro and an R script', *Behavior Research Methods*, 46(2): 331–345.
- Choi, J. J., D. Laibson, B. C. Madrian and A. Metrick (2002), 'Defined contribution pensions: Plan rules, participant choices, and the path of least resistance', *Tax policy and the economy*, 16, 67–113.
- Cohen, J. (1988), *Statistical power analysis for the behavioral sciences*, Hillsdale, NJ: Lawrence Erlbaum Associates.
- Cooper, H. M. (1981), 'On the Significance of Effects and the Effects of Significance', *Journal of Personality and Social Psychology*, 41 (November): 1013–1018.
- Cramér, H. (1946), 'A contribution to the theory of statistical estimation', *Scandinavian Actuarial Journal*, 85–94.
- Crano, W. D., and R. Prislin (2006), 'Attitudes and Persuasion', *Annual Review of Psychology*, 57(1): 345–374.
- Dang, J. (2016), 'Testing the role of glucose in self-control: A meta-analysis', *Appetite*, 107, 222–230.
- Davidai, S., and E. Shafir (2018), 'Are nudges' getting a fair shot? Joint versus separate evaluation', *Behavioural Public Policy*, 1–19.
- Dinner, I., E. J. Johnson, D. G. Goldstein and K. Liu (2011), 'Partitioning Default Effects: Why People Choose not to Choose', *Journal of Experimental Psychology: Applied*, 17(4): 332–341.
- Duval, S., and R. Tweedie (2000), 'Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis', *Biometrics*, 56(2): 455–463.
- Eagly, A. H., and S. Chaiken (1995), Attitude strength, attitude structure, and resistance to change. In R. E. Petty & J. A. Krosnick (Eds.), *Attitude strength: Antecedents and consequences*, Lawrence Erlbaum Associates, Inc.

- * Ebeling, F., and S. Lotz (2015), 'Domestic uptake of green energy promoted by opt-out tariffs', *Nature Climate Change*, 5, 868–871.
- Egger, M., G. Davey Smith, M. Schneider, C. Minder, C. Mulrow, M. Egger, ... I. Olkin (1997), 'Bias in meta-analysis detected by a simple, graphical test', *British Medical Journal*, 315(7109): 629–34.
- * Elkington, J., M. Stevenson, N. Haworth and L. Sharwood (2014), 'Using police crash databases for injury prevention research—a comparison of opt-out and opt-in approaches to study recruitment', *Australian and New Zealand Journal of Public Health*, 38(3): 286–289.
- * Evans, A. M., K. D. Dillon, G. Goldin and J. I. Krueger (2011), 'Trust and self-control: The moderating role of the default', *Judgment and Decision Making*, 6(7): 697–705.
- * Everett, J. A., L. Caviola, G. Kahane, J. Savulescu and N. S. Faber (2015), 'Doing good by doing nothing? The role of social norms in explaining default effects in altruistic contexts', *European Journal of Social Psychology*, 45(2): 230–241.
- Felsen, G., N. Castelo and P. B. Reiner (2013), 'Decisional enhancement and autonomy: public attitudes towards overt and covert nudges', *Judgment and Decision Making*, 8(3): 202–213.
- * Fowlie, M., C. Wolfram, C. A. Spurlock, A. Todd, P. Baylis and P. Cappers (2017), Default effects and follow-on behavior: evidence from an electricity pricing program (No. w23553). National Bureau of Economic Research.
- Franco, A., N. Malhotra and G. Simonovits (2014), 'Publication bias in the social sciences: Unlocking the file drawer', *Science*, 345(6203): 1502–1505.
- Givens, G. H., D. D. Smith and R. L. Tweedie (1997), 'Publication Bias in Meta-Analysis: A Bayesian Data-Augmentation Approach to Account for Issues Exemplified in the Passive Smoking Debate', *Statistical Science*, 12(4): 221–240.
- Griffith, R. W., P. W. Hom and S. Gaertner (2000), 'A Meta-Analysis of Antecedents and Correlates of Employee Turnover: Update, Moderator Tests, and Research Implications for the Next Millennium', *Journal of Management*, 26(3): 463–488.
- Griffiths, L. (2013), Human Transplantation (Wales) Act 2013.
- Hagger, M. S., C. Wood, C. Stiff and N. L. Chatzisarantis (2010), 'Ego depletion and the strength model of self-control: a meta-analysis', *Psychological Bulletin*, 136(4): 495–525.
- * Halpern, S. D., G. Loewenstein, K. G. Volpp, E. Cooney, K. Vranas, C. M. Quill, ... and R. Arnold (2013), 'Default options in advance directives influence how patients set goals for end-of-life care', *Health Affairs*, 32(2): 408–417.
- * Haward, M. F., R. O. Murphy and J. M. Lorenz (2012), 'Default options and neonatal resuscitation decisions', *Journal of Medical Ethics*, 38(12): 713–718.
- * Hedlin, S., and C. R. Sunstein (2016), 'Does active choosing promote green energy use: Experimental evidence', *Ecology LQ*, 43, 107.
- Higgins, J. P. T., S. G. Thompson, J. J. Deeks and D. G. Altman (2003), 'Measuring inconsistency in meta-analyses', *BMJ*, 327(7414): 557–560.
- Ioannidis, J. P. A., J. C. Cappelleri and J. Lau (1998), 'Issues in comparisons between meta-analyses and large trials', *Journal of the American Medical Association*, 279(14): 1089–1093.
- Jachimowicz, J. M. (2017), 'A 5-Step Process to Get More Out of Your Organization's Data', *Harvard Business Review*.
- Jachimowicz, J. M., A. Wihler, E.R. Bailey and A.D. Galinsky (2018), 'Why grit requires perseverance AND passion to positively predict performance', *Proceedings of the National Academy of Sciences*, 115(40): 9980–9985.
- Jachimowicz, J. M., O. P. Hauser, J. D. O'Brien, E. Sherman and A. D. Galinsky (2018), 'The critical role of second-order normative beliefs in predicting energy conservation', *Nature Human Behaviour*, 2(10): 757–764.
- * Jin, L. (2011), 'Improving response rates in web surveys with default setting', *International Journal of Market Research*, 53(1): 75–94

- * Johnson, E. J., and D. Goldstein (2003), 'Do defaults save lives?' *Science*, 302(5649): 1338–1339.
- * Johnson, E. J., S. Bellman and G. L. Lohse (2002), 'Defaults, framing and privacy: Why opting in-opting out', *Marketing Letters*, 13(1): 5–15.
- Johnson, E. J., R. Hassin, T. Baker, A. T. Bajger and G. Treuer (2013), 'Can consumers make affordable care affordable? The value of choice architecture', *PloS One*, 8(12): e81521.
- * E. J. Johnson, J. Hershey, J. Meszaros and H. Kunreuther (1993), 'Framing, probability distortions, and insurance decisions', *Journal of Risk and Uncertainty*, 7(1): 35–51.
- Johnson, E. J., S. B. Shu, B. G. Dellaert, C. Fox, D. G. Goldstein, G. Häubl, ... and E. U. Weber (2012), 'Beyond nudges: Tools of a choice architecture', *Marketing Letters*, 23(2): 487–504.
- Judge, T. A., D. Heller and M. K. Mount (2002), 'Five-factor model of personality and job satisfaction: A meta-analysis', 87(3): 530–541.
- Kahneman, D. (2011), *Thinking, fast and slow*, London, UK: Penguin.
- Kahneman, D., and A. Tversky (1979), 'Prospect Theory: An Analysis of Decision under Risk', *Econometrica*, 47(2): 263–292.
- * Keller, P. A., B. Harlam, G. Loewenstein and K. G. Volpp (2011), 'Enhanced active choice: A new method to motivate behavior change', *Journal of Consumer Psychology*, 21(4): 376–383.
- Kjaergard, L. L., J. Villumsen and C. Gluud (2001), 'Reported Methodologic Quality and Discrepancies between Large and Small Randomized Trials in Meta-Analyses', *Annals of Internal Medicine*, 135(11): 982–989.
- Klein, R. A., K. A. Ratliff, M. Vianello, R. B. Adams, Š. Bahník, M. J. Bernstein, ... B. A. Nosek (2014), 'Investigating variation in replicability: A "many labs" replication project', *Social Psychology*, 45(3): 142–152.
- * Kressel, L. M., and G. B. Chapman (2007), 'The default effect in end-of-life medical treatment preferences', *Medical Decision Making*, 27(3): 299–310.
- * Kressel, L. M., G. B. Chapman and E. Leventhal (2007), 'The influence of default options on the expression of end-of-life treatment preferences in advance directives', *Journal of General Internal Medicine*, 22(7): 1007–1010.
- Krijnen, J. M., D. Tannenbaum and C. R. Fox (2017), 'Choice architecture 2.0: Behavioral policy as an implicit social interaction', *Behavioral Science & Policy*, 3(2): 1–18.
- Kühberger, A. (1998), 'The influence of framing on risky decisions: A meta-analysis', *Organizational Behavior and Human Decision Processes*, 75(1): 23–55.
- La Nación. (2005), El Senado aprobó la ley del donante presunto.
- Landis, J. R., and G. G. Koch (1977), 'The measurement of observer agreement for categorical data', *Biometrics*, 33(1): 159–174.
- Larrick, R. P., and J. B. Soll (2008), 'The MPG illusion', *Science*, 320, 1593–1594.
- Leung, T. (2018), Are you in or are you out? Organ donation in the Netherlands. Retrieved from <https://dutchreview.com/expat/health/system-for-organ-donation-in-the-netherlands/>
- * Li, D., Z. Hawley and K. Schnier (2013), 'Increasing organ donation via changes in the default choice or allocation rule', *Journal of Health Economics*, 32(6): 1117–1129.
- Lipsey, M. W., and D. B. Wilson (2001), *Practical meta-analysis*, Thousand Oaks, CA: Sage.
- * Loeb, K. L., C. Radnitz, K. Keller, M. B. Schwartz, S. Marcus, R. N. Pierson, ... & D. DeLaurentis (2017), 'The application of defaults to optimize parents' health-based choices for children', *Appetite*, 113, 368–375.
- * Löfgren, Å., P. Martinsson, M. Hennlock and T. Sterner (2012), 'Are experienced people affected by a pre-set default option—Results from a field experiment', *Journal of Environmental Economics and Management*, 63(1): 66–72.
- * Madrian, B. C., and D. F. Shea (2001), 'The power of suggestion: Inertia in 401 (k) participation and savings behavior', *The Quarterly Journal of Economics*, 116(4): 1149–1187.
- Marín-Martínez, F., and J. Sánchez-Meca (2010), 'Weighting by Inverse Variance or by Sample Size in Random-Effects Meta-Analysis', *Educational and Psychological Measurement*, 70(1): 56–73.

- McKenzie, C. R. M., M. J. Liersch and S. R. Finkelstein (2006), 'Recommendations implicit in policy defaults', *Psychological Science*, 17(5): 414–420.
- Moshinsky, A., and M. Bar-Hillel (2010), 'Loss Aversion and Status Quo Label Bias', *Social Cognition*, 28(2): 191–204.
- Nakagawa, S., and E. S. A. Santos (2012), 'Methodological issues and advances in biological meta-analysis', *Evolutionary Ecology*, 26(5): 1253–1274.
- * Narula, T., C. Ramprasad, E. N. Ruggs and M. R. Hebl (2014), 'Increasing colonoscopies? A psychological perspective on opting in versus opting out', *Health Psychology*, 33(11): 1426–1429.
- O'Donoghue, T., and M. Rabin (1999), Procrastination in preparing for retirement. In *Behavioral dimensions of retirement economics* (pp. 125–166). Washington, D.C.: Brookings Institution.
- * Or, A., Y. Baruch, S. Tadger and Y. Barak (2014), 'Real-life decision making of Serious Mental Illness patients: Opt-in and opt-out research participation', *The Israel Journal of Psychiatry and Related Sciences*, 51(3): 199–203.
- Peterson, R. A., E. Albaum and R. F. Beltramini (1985), 'A meta-analysis of effect sizes in consumer behavior experiments', *Journal of Consumer Research*, 12(1): 97–103.
- * Pichert, D., and K. V. Katsikopoulos (2008), 'Green defaults: Information presentation and pro-environmental behaviour', *Journal of Environmental Psychology*, 28(1): 63–73.
- * Probst, C. A., V. A. Shaffer and Y. R. Chan (2013), 'The effect of defaults in an electronic health record on laboratory test ordering practices for pediatric patients', *Health Psychology*, 32(9): 995–1002.
- Pustejovsky, J. (2015), clubSandwich: Cluster-robust (sandwich) variance estimators with small-sample corrections.
- * Reiter, P. L., A. L. McRee, J. K. Pepper and N. T. Brewer (2012), 'Default policies and parents' consent for school-located HPV vaccination', *Journal of Behavioral Medicine*, 35(6): 651–657.
- Rosenthal, R. (1995), 'Writing meta-analytic reviews', *Psychological Bulletin*, 118(2): 183–192.
- Sánchez-Meca, J., F. Marín-Martínez and S. Chacón-Moscoso (2003), 'Effect-Size Indices for Dichotomized Outcomes in Meta-Analysis', *Psychological Methods*, 8(4): 448–467.
- Scheibehenne, B., R. Greifeneder and P. M. Todd (2010), 'Can There Ever Be Too Many Options? A Meta-Analytic Review of Choice Overload', *Journal of Consumer Research*, 37(3): 409–425.
- * Shevchenko, Y., B. von Helversen and B. Scheibehenne (2014), 'Change and status quo in decisions with defaults: The effect of incidental emotions depends on the type of default', *Judgment and Decision Making*, 9(3): 287–296.
- Simonsohn, U., L. D. Nelson and J. P. Simmons (2014), 'p-Curve and Effect Size: Correcting for Publication Bias Using Only Significant Results', *Perspectives on Psychological Science*, 9(6): 666–681.
- Slavin, R., and D. Smith (2009), 'The Relationship Between Sample Sizes and Effect Sizes in Systematic Reviews in Education', *Educational Evaluation and Policy Analysis*, 31(4): 500–506.
- * Steffel, M., E. F. Williams and R. Pogacar (2016), 'Ethically deployed defaults: transparency and consumer protection through disclosure and preference articulation', *Journal of Marketing Research*, 53(5): 865–880.
- Sterne, J. A. C., A. J. Sutton, J. P. A. Ioannidis, N. Terrin, D. R. Jones, J. Lau, ... J. P. T. Higgins (2011), 'Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials', *BMJ*, 343(7818).
- Sunstein, C. R. (2015), 'Do People Like Nudges?' *SSRN*.
- Szucs, D., J. P. A. Ioannidis, B. Nosek, G. Alter, G. Banks, D. Borsboom, ... M. Motyl (2015), 'Empirical assessment of published effect sizes and power in the recent cognitive neuroscience and psychology literature', *PLOS Biology*, 15(3).

- Tannenbaum, D., C. Fox and T. Rogers (2017), 'On the misplaced politics of behavioral policy interventions', *Nature Human Behavior*, 1(7): 130.
- Thaler, R. H., and S. Benartzi (2004), 'Save More Tomorrow™: Using Behavioral Economics to Increase Employee Saving', *Journal of Political Economy*, 112(1): 164–187.
- Thaler, R. H., and C. R. Sunstein (2008), 'Nudge: Improving decisions about health, wealth, and happiness', *Penguin Group*.
- * Theotokis, A., and E. Manganari (2015), 'The impact of choice architecture on sustainable consumer behavior: The role of guilt', *Journal of Business Ethics*, 131(2): 423–437.
- Thompson, S. G., and J. Higgins (2002), 'How should meta-regression analyses be undertaken and interpreted?' *Statistics in Medicine*, 21(11): 1559–1573.
- * Trevena, L., L. Irwig and A. Barratt (2006), 'Impact of privacy legislation on the number and characteristics of people who are recruited for research: a randomised controlled trial', *Journal of Medical Ethics*, 32(8): 473–477.
- Trujillo, C. (2013), Uruguay: New Law Renders all Citizens Organ Donors.
- * van Dalen, H. P., and K. Henkens (2014), 'Comparing the effects of defaults in organ donation systems', *Social Science & Medicine*, 106, 137–142.
- Viechtbauer, W. (2010), 'Conducting meta-analyses in R with the metafor package', *Journal of Statistical Software*, 36(3): 1–48.
- Willsher, K. (2017), France introduces opt-out policy on organ donation. Retrieved from <https://www.theguardian.com/society/2017/jan/02/france-organ-donation-law>
- * Young, S. D., B. Monin and D. Owens (2009), 'Opt-Out Testing for Stigmatized Diseases: A Social Psychological Approach to Understanding the Potential Effect of Recommendations for Routine HIV Testing', *Health Psychol*, 28(6): 675–681.
- * Zarghamee, H. S., K. D. Messer, J. R. Fooks, W. D. Schulze, S. Wu and J. Yan (2017), 'Nudging charitable giving: Three field experiments', *Journal of Behavioral and Experimental Economics*, 66, 137–149.
- Zlatev, J. J., D. P. Daniels, H. Kim and M. A. Neale (2017), 'Default neglect in attempts at social influence', *Proceedings of the National Academy of Sciences*, 114(52): 13643–13648.
- Zúñiga-Fajuri, A. (2015), 'Increasing organ donation by presumed consent and allocation priority: Chile', *Bulletin of the World Health Organization*, 93(3): 199–202.

Appendix A. Default meta-analysis coding scheme

The following coding scheme was developed to investigate possible underlying channels of default effects in existing studies. A channel is a pathway through which the effects of making one choice option the default can occur. For example, a default's effectiveness may unfold through the endorsement that is implied by the default; namely, decision-makers may believe that the choice of default suggests which course of action is recommended by the choice architect.

Three channels for defaults' effectiveness have been identified in the prior literature, and the effect of any given default in a study may happen through three, two, one or none of these channels. Presumably, a default has a stronger effect if more channels are involved. The default effect may also vary depending on how strongly each channel is involved.

The aim of this coding scheme is for you to provide expert judgment of whether and how strongly each of the three channels should be expected to be involved in each of the default studies that we have identified.

Below, we outline what each of the three channels is. We hope that, in the end, you will be able to provide three scores for each default study, describing the extent to which you think each of the three channels is involved in that study (i.e., one score for each channel). Obviously, this is a subjective assessment, but your training and your personal introspection will hopefully allow you to make this type of assessment.

Endorsement

The decision-maker perceives the default as conveying what the choice architect thinks the decision-maker should do. For example, setting organ donation as the default communicates to decision-makers what the choice architect believes is the 'right' thing to do. One factor that may therefore influence how much this channel will influence a default's effectiveness is how much the decision-makers trust and respect the architect of the decision-making design.

For this rating/code, we would like you to rate the extent to which you think decision-makers perceived the default as a favorable recommendation from the choice architect. The scale has three levels: '0' (this channel does not play a role), '1' (this channels plays somewhat of a role) and '2' (this channel plays a role).

Ease

Defaults are effective in part because it is easier for individuals to stay with the pre-selected option than to choose a different option. The decision of whether or not to stay with the default may then be influenced by how difficult it is to change the default. For example, if it is particularly difficult to opt out of a default (i.e., when the steps that one has to take in order to switch the default require a lot of effort), then ease may underlie the default's effectiveness. The more effort it takes to change the default, the more likely individuals may be to stay with the pre-selected option.

For this rating/code, we would like you to rate how difficult you think changing the default option was. The scale has three levels: '0' (this channel does not play a role), '1' (this channels plays somewhat of a role) and '2' (this channel plays a role).

Endowment

The effectiveness of a default also varies depending on the extent to which decision-makers think about the pre-selected option as the status quo. The more decision-makers feel endowed with the pre-selected option and evaluate other options in comparison to it, the more likely they are to stay with the default. For example, if the default has been in place for a while and therefore has been part of the decision-maker's life, then they are likely to feel more endowed with it. Endowment with the default may be greater when the default is presented in a way that reinforces the belief that the default is the status quo. Endowment with the default may also be greater when the decision-maker has little experience in the choice domain.

For this rating/code, we would like you to rate how much you think decision-makers felt endowed with the default option. The scale is binary: '0' (this channel does not play a role) or '1' (this channel plays a role).

Appendix B. Default meta-analysis channel scatter plot and correlations

