Default Behavior and Social Inference

Abstract

Default options can convey information relevant to making a default decision. We find that automatic enrollment defaults (where individuals *opt out* of a program) are often thought to convey information about what course of action is sensible or best, whereas nonenrollment defaults (individuals *opt into* a program) are often viewed as uninformative about a choice architect's intentions and beliefs. This "recommendation asymmetry" arises because of a difference in the reasons imputed to the choice architect when choosing automatic enrollment versus nonenrollment. Individuals assume automatic enrollment reflects a deliberate decision by the choice architect who selected the default for goal-directed reasons, whereas nonenrollment often reflects simple inaction without any supporting reasons. In both laboratory and field experiments we demonstrate that individuals draw reliably different inferences from automatic and nonenrollment defaults, and use this information when making a default decision. Furthermore, we test and corroborate an important implication of this hypothesis, namely that feelings of trust in the choice architect will have be more predictive of default acceptance for automatic enrollment than nonenrollment. Our findings have implications for both policy design and libertarian paternalism.

Keywords— Choice architecture, Nudges, Defaults, Social Inference, Trust

Introduction

Much of public policy is predicated upon individual choice. Individuals decide what foods to eat, which medical procedures to undertake, what types of financial investments to make, and how much to save for retirement. But people do not always actively decide, even in instances where a decision must be made. In such cases a third party often sets a *default option*, a designated course of action for those who fail to explicitly choose for themselves.

Interest in defaults has exploded in recent years because of their practical importance. Compared to a nonenrollment default, governments that presume citizens as willing organ donors have markedly higher donation rates (Abadie and Gay, 2006; Johnson and Goldstein, 2003); employees at companies with automatic 401(k) enrollment defaults save more for retirement (Choi et al., 2002; Madrian and Shea, 2001); cities with "green" electricity defaults result in more efficient energy use (Pichert and Katsikopoulos, 2008); and states with limited tort defaults have drivers who pay lower insurance premiums (Johnson et al., 1993). Default effects have also been observed in the use of advanced medical directives, internet privacy preferences, legal contracts, medical vaccine adherence, and colorectal cancer screenings (Bellman et al., 2001; Chapman et al., 2010; Johnson et al., 2002; Kressel et al., 2007; Korobkin, 1998; Mehta et al., 2018; Young et al., 2009). The strategic selection of default options is an attractive policy tool because it is often a cost-effective alternative to traditional policy approaches (Benartzi et al., 2017), and does so by "nudging" behavior without meaningfully limiting freedom of choice among a set of options (Thaler and Sunstein, 2008; Sunstein, 2013).

Defaults are also of theoretical interest because classic models of choice predict that — contrary to the evidence — their effects should be negligible. Since defaults do not change the set of options available to individuals, and the costs of explicitly registering a preference are usually trivial, they should have little to no influence on any individual with a stable preference ordering over options. However, past research finds that defaults are effective for at least three reasons: (i) because people display inertia, (ii) because defaults serve as reference points towards which people tend to be biased, and (iii) because defaults can be viewed as an implicit recommendation by the choice architect (Dinner et al., 2011; Johnson and Goldstein, 2003; McKenzie et al., 2006).

In this paper we focus on the last of these three mechanisms. While previous work has demonstrated that individuals can view default options as an implicit recommendation and that such inferences affect default decisions (McKenzie et al., 2006), little work has examined the conditions that give rise to this process or the downstream consequences of such inferences on the effectiveness of default policies. Using laboratory and field experiments, we demonstrate that individuals reliably view some types of defaults as a clearer signal of implicit advice by the choice architect, and use this information when making a default decision. In particular, people tend to view automatic enrollment defaults (where individuals *opt out* of a program) as especially diagnostic of what the choice architect thinks is best or sensible, whereas nonenrollment defaults (individuals *opt into* a program) are seen as relatively uninformative about a choice architect's intentions and beliefs. This asymmetry, we argue, is due to a fairly sophisticated set of inferences that participants draw from the construction of a default option, and has implications for when policy defaults will be more or less effective.

Inferring Reasons from Defaults

In this paper we make a distinction between automatic enrollment and nonenrollment defaults. We operationalize automatic enrollment as any default that alters an individual's state of affairs when they fail to explicitly choose for themselves, whereas nonenrollment defaults do not alter an individual's state of affairs in the absence of choice. As an illustration, consider employer 401(k) plans in the United States. When an employee fails to explicitly choose how to save for retirement, they can either be defaulted into a savings plan (and by doing so, their state of affairs is changed to one where they are now saving with the company) or be defaulted into no savings plan (which was what the employee was already doing before joining the company). Although not all choice menus with defaults can

be partitioned according to this distinction between automatic enrollment and nonenrollment, many can. Our key hypothesis is that automatic enrollment defaults are more likely than nonenrollment defaults to send a signal of implicit advice from the choice architect. Upon receiving this signal of advice, individuals who trust the choice architect will be more likely to stay with the default option whereas individuals who distrust the choice architect will be less likely to stay with the default.

To motivate this hypothesis, we start with the working assumption that people explain organizational policies the same way they explain most behavior — by appealing to an agent's reasons and intentions. Referencing an agent's subjective reasons for acting is the predominant way that people explain both their own behavior and that of others (Baker et al., 2009; Heider, 1958; Malle, 2004; Shafir et al., 1993). Furthermore, decision makers recognize that others search for reasons when making sense of their decisions and factor such attributional implications into the decision making process (Medin et al., 1999). Because policy decisions are especially consequential, it is likely that people expect choice architects to have cogent reasons for their policy decisions and that choice architects would understand this.

More to the point, people may assume that some defaults are only implemented when there is a sufficiently strong justification for doing so, and this assumption likely holds for automatic enrollment. Only automatic enrollment defaults have the potential to alter an individual's state of affairs in the absence of choice, and actions that alter the status-quo require stronger justifications than those that do not (Spranca et al., 1991; Baron and Ritov, 1994). This reluctance to engage in action is amplified when choice architects are held accountable for their choices (Tetlock and Boettger, 1994) and when policy decisions are made by committee (White et al., 2011), two often-present conditions when implementing policies within organizations. For these reasons, individuals may tacitly understand that choice architects would be unlikely to select automatic enrollment defaults without strong justifications for doing so.[1]

By contrast, nonenrollment defaults do not alter an individual's state of affairs when they fail to choose for themselves, so enacting them should require less compelling reasons. They may be in place because choice architects perceive non-action as the norm (Thaler and Sunstein, 2003); because choice architects are uncertain about what option is best and so stick with doing nothing (Tetlock and Boettger, 1994); because of accountability concerns for changing an individual's state of affairs (DeScioli et al., 2011); or because choice architects never made an explicit decision in the first place (Kordes-de Vaal, 1996). In the absence of strong reasons for choosing one default over another, nonenrollment is likely to be viewed as the default option for setting a default option. And since nonenrollment defaults lend themselves to multiple explanations, they are less likely to convey information about the choice architect's preferences (Morris and Larrick, 1995).

In short, there is a unique plausible reason for why choice architects would select an automatic enrollment default (to encourage enrollment), but multiple plausible reasons for why a choice architect would select a nonenrollment default.[2] The recommendation asymmetry in default options naturally follows from the difference in reasons imputed to the choice architect when choosing automatic enrollment versus nonenrollment. If individuals assume that choice architects primarily select automatic enrollment defaults when they have strong justifications for doing so, then people

---

[1]This contention is also consistent with United States legislation on retirement savings policy. In 2006 the U.S. government passed the Pension Protection Act (PPA), considered one of the largest reforms to pension legislation in decades (see Beshears et al., 2010c). The PPA was partly designed to encourage employers to structure defined contribution saving plans more aggressively through the use of automatic enrollment defaults, and this enticement was done by reducing legal liability for potential losses resulting from such plans (given that other requirements were also met, such as exposure to more than one asset class). This enticement assumes that employers are not selecting automatic enrollment saving plans partly because of risk-aversion, which is consistent with the behavioral evidence reviewed here.

[2]Formally, we can think of the recommendation asymmetry as a kind of *information bottleneck*, where a policy default acts as a signal that is more or less efficient in conveying information about states of the world (Skyrms, 2009). As a simple example, think of a conceptual state space where choice architects find themselves in one of three worlds: (i) they have reason to encourage enrollment, (ii) they have reason to not encourage enrollment, or (iii) they are indifferent. When there are more possible states of the world than default options, some defaults will more finely partition the state space than others and in doing so will send more a informative signal about the state space. This analysis allows us to quantify the amount of information (in bits) provided by a default option using a conventional measure of information gain (KL divergence, or D_{KL}; Kullback and Leibler, 1951):

will view the default as an implicit recommendation as a matter of course. Defaults that are implemented with no particularly compelling reason in mind, on the other hand, simply do not speak to what a choice architect endorses or recommends.

The most direct evidence for a recommendation asymmetry comes from a set of experiments by McKenzie et al. (2006). In one study, participants were placed in the role of "policymaker" and tasked with selecting an organ donation default. Participants who selected an automatic enrollment default were largely comprised of individuals who explicitly stated that citizens should be organ donors (81%). By contrast, participants who selected the nonenrollment default were split between those who did not think citizens should be donors (31%), those who thought citizens should be donors (48%), and those who were unsure what to think (21%). In this way, participants who selected an automatic enrollment default revealed more information about their preferences than did participants who selected a nonenrollment default. In a follow-up study, participants were asked to draw an inference from the default option for organ donation. Consistent with a recommendation asymmetry, participants were especially likely to view the automatic enrollment as a recommendation by the choice architect (see McKenzie and Nelson, 2003, for a similar finding). While consistent with our account, the results of McKenzie et al. do not provide conclusive evidence for a recommendation asymmetry because it is unclear whether participants make such inferences conditional upon observing a default option, or would also draw a similar set of inferences in the absence of a default option. If participants start with the prior belief that most people have a preference towards organ donation, then it could be the case that the default option does not provide additional information or even that *nonerollment* is the one that communicates unique information about a choice architect's preferences. Thus, to directly establish a recommendation asymmetry in default option requires accounting for prior beliefs about what is recommended or preferred. We discuss this issue in more detail in Study 1, and provide a direct test of the recommendation asymmetry that controls for prior beliefs.

Implicit Recommendations and Trust in the Choice Architect

What are the implications of the recommendation asymmetry on default behavior? If people infer that a default was selected by the choice architect for goal-directed reasons, then it is natural to ask whether those reasons are ones that benefit only the choice architect. How inferences of implicit advice will affect decisions to stay with a default option will likely be qualified by whether the choice architect is thought to be trustworthy or untrustworthy.

To illustrate these points, consider two individuals Barb and Jimmy. Barb observes a choice architect who selects an automatic enrollment default and Jimmy observes a choice architect who selects a nonenrollment default. Barb and Jimmy are equally distrustful of the choice architect, and assume the choice architect's interests are negatively correlated with their own. In our model, Barb will still view automatic enrollment as informative ("The choice architect must have set this as the default for a particular reason . . . "), but she draws a negative conclusion from that information (". . . which is to separate me from my wallet") that makes her *less* likely to stay with the default option than if she was trusting of the choice architect (also see Friestad and Wright, 1994). By contrast, Jimmy views the nonenrollment default as not particularly informative about the choice architect's reasons and intentions, and so his feelings of distrust do not figure as prominently into his decision to stay with or leave the default option. Because automatic enrollment defaults are predicted to be an especially strong signal of implicit advice by the choice architect, default decisions should be more

$$D_{KL}(default) = \sum_{i=1}^{n} p(state_i \mid default) \log_2 \left[\frac{p(state_i \mid default)}{p(state_i)} \right] \tag{1}$$

where $state_i$ represents a particular state of the world. The information provided by a default option is the total reduction in uncertainty (compared to prior beliefs) upon learning that a choice architect selected a particular policy default. We can think of the recommendation asymmetry as the information acquired over a subset of states, namely those that signal a choice architect's preference towards enrollment or nonenrollment. As we discuss in Study 1, to measure prior beliefs (i.e., the denominator in the right-hand term of equation 1) we compare inferences from automatic enrollment and nonenrollment defaults to inferences made using a no-default control condition.

sensitive to one's feelings of trust in the choice architect under automatic enrollment compared to nonenrollment.

The recommendation asymmetry allows us to make a novel prediction about how trust in the choice architect will impact default behavior. If the basic recommendation asymmetry holds, then we can expect that trust should be more predictive of default behavior for automatic enrollment than for nonenrollment defaults. For automatic enrollment, we should expect to see pronounced default behavior when trust is high and modest default behavior (perhaps even anti-default behavior) when trust is low; with nonenrollment defaults, we should be expect a less dramatic swing because in this case the default communicates less information about what the choice architect thinks.

Overview of Studies

We hypothesize that automatic enrollment defaults are more likely than nonenrollment defaults to signal an implicit recommendation by the choice architect. In Study 1 we directly test this asymmetry across a number of policy settings by comparing inferences from automatic and nonenrollment to a control condition. In Study 2 we examine a potential explanation of the recommendation asymmetry, namely that automatic enrollment defaults are viewed as more paternalistic than nonenrollment defaults and so demand stronger justifications for being implemented. In Study 3 we test an important implication of the recommendation asymmetry, namely that feelings of trust in choice architect will be more predictive of default decisions for automatic enrollment than nonenrollment defaults. Finally, in Study 4 we examine the link between default inferences, trust, and behavior in a field experiment with real stakes.

For each study, we set our sample size before data were collected. All research was approved by either the University of California, Irvine, or the University of Utah Institutional Review Boards. The Supplementary Materials includes study materials, data, and code for all studies.

Study 1

We begin by testing the hypothesis that automatic enrollment defaults are more likely than nonenrollment defaults to signal an implicit recommendation by the choice architect. One challenge of testing for a recommendation asymmetry is that participants may hold prior beliefs over what options are recommended. For instance, in McKenzie et al. (2006) participants were more likely to view a retirement savings default as the recommended option when employees were automatically enrolled into a savings plan (and could opt-out), compared to when employees were not enrolled into a savings plan (but could opt-in). Although this result is consistent with a recommendation asymmetry, it could simply reflect the more mundane fact that many people recommend saving for retirement, and few people recommend the opposite. For this reason, comparing inferences from automatic and nonenrollment defaults is not sufficient to demonstrate that participants acquire new information from the default option about what the choice architect recommends.

To properly test for a recommendation asymmetry, one useful approach is to compare responses to a baseline condition. Doing so allows for a comparison to inferences that participants draw in the absence of a default option. *Mandated choice* policies, where no default option exists and individuals must instead explicitly state their preferences, serve as a natural baseline comparison for such a task. In Study 1, we test inferences made under automatic enrollment and nonenrollment defaults to inferences made under mandated choice across a range of policy settings. We pre-registered the design and analysis for Study 1 at `http://aspredicted.org/blind.php?x=9nf76y`.

Method

We recruited a sample of 451 U.S. adults from Amazon.com's Mechanical Turk online labor market (MTurk) to participate in return for a flat cash payment (48% female, mean age[3] = 37 years, age range: 18–74 years). Participants were required to complete the study using a non-mobile device. Before completing the study, participants responded to a simple attention check, and only those who passed the attention check were allowed to participate in the study. Participants were randomly assigned to conditions after passing the attention check.

Our study used a 4 (policy setting) × 3 (policy regime) mixed factorial design, with policy setting randomized within-participants and policy regime randomized between-participants. Participants responded to four policy descriptions, each presented on separate pages and in random order: (i) a company retirement savings policy for employees; (ii) a government organ donation policy for citizens; (iii) a hospital medical testing policy for patients; and (iv) a bank overdraft protection policy for customers. We chose these policy settings because they cover both government and business settings, and because the target behavior is presumably familiar and easy for participants to understand. For each description, participants were randomly assigned to either (i) a mandated choice policy regime, (ii) an automatic enrollment default regime, or (iii) nonenrollment default regime. As an example, the three retirement savings policy descriptions were as follows:

> *Mandated choice*:
> Imagine you have recently taken a job at a new company. This company has the following policy for employee retirement savings: **New employees are required to indicate whether or not they wish to enroll into a retirement savings plan.**

> *Nonenrollment default*:
> Imagine you have recently taken a job at a new company. This company has the following policy for employee retirement savings: **New employees are only enrolled into a retirement savings plan when they explicitly request to participate in a savings plan.**

> *Automatic enrollment default*:
> Imagine you have recently taken a job at a new company. This company has the following policy for employee retirement savings: **New employees are automatically enrolled into a retirement savings plan unless they explicitly request not to participate in a savings plan.**

For each policy description, participants were asked two questions adapted from McKenzie et al. (2006). The first question asked about implicit advice, and took the form of "What, if anything, does this policy say about what policymakers think their constituents ought to do?" They selected one of three options, indicating that policymakers probably think their constituents ought to be enrolled, that policymakers probably do not think their constituents ought to be enrolled, or that "this policy tells me nothing" about what policymakers think their constituents ought to do. The second question asked about majority preference, and took the form of "What, if anything, does this policy say about what policymakers think most of their constituents prefer?" The response options were that policymakers probably think most constituents prefer to be enrolled, that policymakers probably do not think most constituents prefer to be enrolled, or that the policy says nothing about what policymakers think most constituents prefer. All study materials and questions can be found in the Supplemental Material.

---

[3]Two participants reported their age as 0, which we do not include when reporting age statistics.

Table 1: Response Percentages in Study 1

	Retirement Savings			Overdraft Protection			Organ Donation			Medical Testing		
	MC	NE	AE	MC	NE	AE	MC	NE	AE	MC	NE	AE
Policymakers think constituents ought to be enrolled?												
Yes	29.5	24.2	85.4	30.7	13.9	78.4	23.1	20.0	85.0	31.8	10.6	81.0
No	9.4	17.4	6.0	14.3	41.8	8.5	9.6	15.5	3.6	14.6	58.2	7.0
Can Infer Nothing	61.2	58.4	8.6	55.0	44.3	13.1	67.3	64.5	11.4	53.6	31.2	12.0
Policymakers think most constituents prefer to be enrolled?												
Yes	25.2	22.4	66.9	26.4	20.3	71.2	11.5	9.0	45.7	18.0	10.6	44.9
No	11.5	36.0	15.9	21.4	41.8	5.9	19.9	43.2	22.1	21.3	48.9	20.9
Can Infer Nothing	63.3	41.6	17.2	52.1	38.0	22.9	68.6	47.7	32.1	60.7	40.4	34.2

Notes: Table numbers reflect column-wise percentages for each question. MC = mandated choice policy, NE = nonenrollment default, AE = Automatic enrollment default. The first question asked about implicit advice, and took the form of "What, if anything, does this policy say about what policymakers think their constituents ought to do?" They selected one of three options, indicating that policymakers probably think their constituents ought to be enrolled, that policymakers probably do not think their constituents ought to be enrolled, or that "this policy tells me nothing" about what policymakers think their constituents ought to do. The second question asked about majority preference, and took the form of "What, if anything, does this policy say about what policymakers think most of their constituents prefer?" The response options were that policymakers probably think most constituents prefer to be enrolled, that policymakers probably do not think most constituents prefer to be enrolled, or that the policy says me nothing about what policymakers think most constituents prefer. Cell numbers may not sum exactly to 100 due to rounding error.

Results

We test for a recommendation asymmetry by comparing inferences in each default condition to that of mandated choice. Displayed in Table 1 are the response proportions for each question, separated by policy condition and policy setting. Looking first at inferences of implicit advice, for mandated choice the modal response in every setting is that the policy says nothing about what the choice architect thinks their constituents ought to do. We see a roughly similar pattern for nonenrollment defaults, where the modal response was also that the policy "says nothing" in three of the four settings. By contrast, for automatic enrollment, the majority of participants in every policy setting inferred that choice architects thought participants ought to be enrolled in the default option. Compared to our baseline condition, participants appear more likely to shift their inferences after viewing an automatic enrollment default than after viewing a nonenrollment default.[4] Table 1 shows a similar, albeit weaker, pattern for inferences of majority preference.

To more precisely test for a recommendation asymmetry, we compared the likelihood of drawing a default-consistent inference for each policy default, relative to mandated choice. For example, we compared the likelihood of reporting that policymakers think constituents should be enrolled under automatic enrollment, compared to the same response under mandated choice; for nonenrollment, we compared the likelihood of reporting that policymakers did not think constituents should be enrolled, compared to the same response under mandated choice. Table 2 displays the marginal increase (i.e., percentage point difference) in responding over baseline for automatic enrollment and nonenrollment across policy settings. The "asymmetry score" columns report the difference between the two marginal effects, with

---

[4]To quantify the overall difference in responding across conditions, for each question and policy setting we calculated Cramer's ϕ — a measure of association when using nominal variables — between mandated choice and automatic enrollment, and between mandated choice and nonenrollment. For all eight comparisons (four policy settings with two questions each), we find a greater shift in responding (i.e., higher ϕ values) for automatic enrollment than for nonenrollment. To test for differences in ϕ between automatic and nonenrollment, we aggregated over policy settings and constructed bootstrapped standard errors of ϕ for each comparison (based on 10,000 resamples, clustered at the participant-level). Comparing the two bootstrapped estimates, for implicit advice we find a relatively larger shift in responding (over baseline) for automatic enrollment than nonenrollment, $Z = 7.49$, $p < 0.001$. We observe a similar pattern for inferences of majority preference, $Z = 3.65$, $p < 0.001$.

Table 2: Recommendation Asymmetry Results in Study 1

	Policymakers recommend default?			Policymakers think most prefer default?		
	$p(\text{Yes}|\text{AE})-$ $p(\text{Yes}|\text{MC})$	$p(\text{No}|\text{NE})-$ $p(\text{No}|\text{MC})$	asymmetry score	$p(\text{Yes}|\text{AE})-$ $p(\text{Yes}|\text{MC})$	$p(\text{No}|\text{NE})-$ $p(\text{No}|\text{MC})$	asymmetry score
Retirement savings	55.93 (4.82)	8.04 (3.88)	47.90*** (6.50)	41.71 (5.32)	24.51 (4.66)	17.19[†] (7.36)
Overdraft protection	47.72 (5.13)	27.49 (4.92)	20.23** (7.54)	44.81 (5.23)	20.34 (5.24)	24.47** (7.93)
Organ donation	61.92 (4.53)	5.87 (3.75)	56.05*** (6.12)	34.18 (4.93)	23.35 (5.11)	10.82* (7.31)
Medical testing	49.22 (4.91)	43.59 (5.06)	5.64 (7.47)	26.94 (5.06)	27.60 (5.38)	−0.67 (7.72)
All results combined	53.71 (2.77)	21.00 (2.29)	32.71*** (3.62)	36.82 (3.01)	23.93 (2.64)	12.89*** (4.13)

Notes: Table numbers reflect the marginal change (i.e., percentage point difference) in response options between each default condition and mandated choice, with robust standard errors in parentheses. MC = mandated choice policy, NE = nonenrollment default, AE = Automatic enrollment default. The first column compares the likelihood of inferring that policymakers recommend enrollment under an automatic enrollment default compared to mandated choice. The second column compares the likelihood of inferring that policymakers do not recommend enrollment under a nonenrollment default compared to mandated choice. The third column ("asymmetry score") takes the difference between the first two columns, with positive numbers indicating a recommendation asymmetry in favor of automatic enrollment. The fourth column compares the likelihood of inferring that policymakers think most constituents prefer enrollment under an automatic enrollment default compared to mandated choice. The fifth column compares the likelihood of inferring that policymakers do not think most constituents prefer enrollment under a nonenrollment default compared to mandated choice. The sixth column computes an asymmetry score similar to that in the third column. For marginal effects based on the aggregate model (i.e., "All results combined" row), robust standard errors are clustered by participants. *P*-values are calculated from the multinomial logit model described in Footnote 5. [†]$p < 0.10$, *$p < 0.05$, **$p < 0.01$, *** $p < 0.001$

positive values indicating a larger shift under automatic enrollment than nonenrollment.[5]

First looking at inferences of implicit advice, we see that in every policy setting that asymmetry score is positive — indicating that automatic enrollment sent a stronger signal of implicit advice than did nonenrollment — and significant at $p < 0.05$ in three of the four settings. Across policy settings, when viewing a nonenrollment default participants were on average 21 percentage points more likely (over baseline) to infer that choice architects did not recommend enrollment. By contrast, when viewing an automatic enrollment default participants were on average 54 percentage points more likely (over baseline) to infer that choice architects recommended enrollment. Consistent with the recommendation asymmetry hypothesis, the treatment effect for automatic enrollment was reliable larger than that observed for nonenrollment, $Z = 6.44$, $p < 0.001$. We find a qualitatively similar, albeit weaker, pattern of results when examining inferences of majority preference. The asymmetry score is positive — indicating that automatic enrollment sent a stronger signal of majority preference than did nonenrollment — in three of four policy settings (for the medical testing scenario, the asymmetry score was essentially zero). Across policy settings, when viewing a nonenrollment default participants were on average 24 percentage points more likely (over baseline) to infer that choice architects probably did not think constituents preferred enrollment. By contrast, when viewing an automatic enrollment default participants were on average 37 percentage points more likely (over baseline) to infer that choice architects probably thought constituents preferred enrollment. As expected, the treatment effect for automatic enrollment was reliably

---

[5]To estimate standard errors and calculate *p*-values, we conducted a multinomial logit regression for each question and policy setting, with the "policy tells me nothing" response option serving as our reference outcome value. Each model included dummy variables for automatic enrollment and nonenrollment, with mandated choice serving as our reference condition. We also implemented robust standard errors for each model. Since mandated choice was used as our reference condition, we were able to estimate the marginal effect of choosing a particular response option (over baseline) for each policy default. To calculate *p*-values for asymmetry scores, we conducted an equality test of coefficients between the default-consistent response for each policy default. When aggregating over policy settings, the multinomial logit models also included policy setting fixed effects and participant-clustered robust standard errors.

larger than that observed for nonenrollment, $Z = 3.54$, $p < 0.001$.

The results of Study 1 corroborate the hypothesis that automatic enrollment defaults are especially likely to signal an implicit recommendation by the choice architect. Compared to our control condition, participants were more likely to shift their inferences about implicit advice and majority preference after viewing an automatic enrollment default than after viewing a nonenrollment default. It is also worth mentioning that although nonenrollment sent a relatively weaker signal, compared to our control condition these policies did nevertheless communicate a (negative) signal of discouraging enrollment. These results suggest that prior beliefs do appear to play a role in the data, but are not sufficient to explain the recommendation asymmetry.

Follow-up Studies

In the Supplemental Materials we report the results of three additional studies (Studies S1-S3) which replicate and extend the scope of the recommendation asymmetry. Our first supplemental study (Study S1) replicates the findings of Study 1 using an alternative response format. Unlike Study 1, where participants chose from one of three fixed response options, in Study S1 ($N = 150$) we used a continuous response scale that allowed respondents to provide gradations in the likelihood of each inference, and also did not force participants to choose between competing inferences. We find that the recommendation asymmetry is robust to this alternative elicitation procedure. Participants were again more likely to infer inferences of implicit advice and majority preference from an automatic enrollment default than from a nonenrollment default.

In our second study (Study S2), we take an alternative approach to ruling out prior beliefs as an alternative explanation to the recommendation asymmetry. To do so, we exploited the granularity required by employers when selecting a default retirement savings plan for employees. We presented a new sample of MTurkers ($N = 400$) with descriptions of one of four retirement savings policies. One version was a nonenrollment default (employees only saved for retirement when they explicitly elected into a savings plan), whereas the other three versions involved an automatic enrollment policy that varied the default savings rates at either 3%, 5%, or 7% of an employee's earnings. If participants are simply applying existing knowledge, and not acquiring specific information from the selection of the default option, then responses across the three variations of automatic enrollment should be the same. We find however that responses vary systematically across the three automatic enrollment conditions, with a majority of participants viewing their particular default allocation as a reflection of what the choice architect thought was optimal for employees, best for employees, and a signal of majority preference (all p-values < 0.001). Furthermore, and consistent with Study 1, participants were more likely to view the default as uninformative in the nonenrollment condition than in the automatic enrollment conditions.

In our third study (Study S3) we examined an alternative explanation for the recommendation asymmetry, namely that participants draw different inferences from automatic enrollment and nonenrollment defaults based on the frequency with which they are implemented. From a Bayesian perspective rare events are more informative than common events (e.g., McKenzie and Mikkelsen, 2007), and so participants may draw stronger inferences from automatic enrollment than nonenrollment because the former is thought to be less frequent. A related argument is that commonplace policies usually serve as the status-quo, and it is deviations from the status-quo (in the form of an infrequent policy) rather than automatic enrollment that signal information about what option is recommended or preferred. This account would predict that manipulating policy frequency independent of the policy default should eliminate the recommendation asymmetry. To test this account, we presented a new sample of MTurkers ($N = 162$) with policy descriptions similar to Study 1, but this time only for automatic and nonenrollment default policies. We also independently varied whether the policy default was common or uncommon by providing participants with statistical information about how frequent enrollment behavior was. The results of Study S3 suggest the recommendation asymmetry cannot be explained by a

frequency-based account. Across policy settings participants were more likely to infer implicit advice from an automatic enrollment default than a nonenrollment default (58% marginal effect for automatic enrollment over nonenrollment, $p < 0.001$). However, whether the default was statistically common or uncommon had little effect on such inferences (-0.1% marginal effect for uncommon defaults, $p = 0.982$), nor did we observe a reliable interaction between frequency and policy default ($p = 0.522$ for the interaction term). In fact, we find on average that participants viewed a *common* automatic enrollment default as a stronger signal of implicit advice than a *rare* nonenrollment default, a finding directly at odds with a frequency-based explanation of the recommendation asymmetry.

Study 2

In Study 1, we document a recommendation asymmetry in favor of automatic enrollment defaults. In Study 2 we examine one possible explanation for this asymmetry, namely that in the absence of supporting reasons people assume that choice architects will tend to favor nonenrollment defaults. This occurs, we argue, because automatic enrollment defaults alters an individual's state of affairs when they fail to choose for themselves. In this sense an automatic enrollment default can be seen as a more interventionist or paternalistic policy than nonenrollment, even though the two policies contain the exact same set of available options and are therefore formally equivalent. Because people assume that more paternalistic actions require greater justification than less interventionist ones (e.g., Lupoli et al., 2018), they take the existence of an automatic enrollment default as evidence for such a justification ("the choice architect must have had a good reason for choosing automatic enrollment, otherwise they would not have selected this option as the default"). Study 2 tests this hypothesis by measuring increases or decreases in judgments of "justifiability" under policies that vary in their degree of rated paternalism.

In Study 2 we compare automatic enrollment and nonenrollment defaults to *mandated choice* and *mandated enrollment* policy regimes, which represent opposite ends of the paternalism spectrum. As discussed in the introduction to Study 1, mandated choice policies do not implement a default option and individuals must instead explicitly state their preferences. Mandated enrollment policies, on the other hand, do not directly consult an individual's preferences at all and instead policymakers select a course of action by fiat or formal decree. For instance, several U.S. states have "routine removal" policy in which they can remove an individual's corneas upon their death without permission from the deceased's next of kin.[6] Since individuals are not given a choice in what course of action to pursue, mandated enrollment is more paternalistic than both mandated choice and any kind of policy default.

If implementing paternalistic policies requires stronger reasons and justifications than implementing less paternalistic policies, then we should expect mandated choice to score especially low, and mandated enrollment to score especially high, on both measures of paternalism and "justifiability." Default options are expected to fall somewhere in between, with automatic enrollment scoring higher than nonenrollment. This pattern of findings would suggest a close tie between paternalism and signals of tacit advice by the choice architect (because that policy requires strong justifications for its implementation). However, if inferences of implicit advice are tracking other features of policy defaults — such as the effort costs associated with leaving a default option or whether individuals have the same set of options available to choose from — then we should not expect systematic differences in judgments between automatic and nonenrollment defaults.

---

[6]In practice, routine removal policies do allow for opting-out if individuals or family members object, but unlike a presumed consent default (i.e., automatic enrollment) there is no obligation for medical practitioners to notify the next of kin. Since most citizens are unfamiliar with routine removal policies, these policies in effect provide little in the way of individual choice. More abstractly, the crucial difference between automatic enrollment and routine removal policies is whether the rights of a deceased body lay with the state or the individual and their next of kin (for a discussion, see Childress and Liverman, 2006).

Method

We recruited a sample of 201 U.S. participants from MTurk (62% female, average age = 33 years, range: 18–68 years). We excluded one additional participant who reported an age less than 18 years old; including this participant in the analysis does not change the direction or significance levels of any of our results.

All participants were asked to make comparative judgments of two different state policies regarding human cornea donation. Cornea donation was chosen because policies vary considerably from state to state (Lee et al., 1989), and participants were presumed to have few preconceptions about such policies.[7] Participants first read a description about the general purpose of cornea donations, and were then provided with policies from two different states (generically labeled State A and State B). Each state policy was described in one of four ways: (i) *mandated choice*: citizens were required to explicitly state their preference for cornea donation; (ii) *nonenrollment*: citizens were presumed to be non-donors unless they explicitly opted into donation; (iii) *automatic enrollment*: citizens were presumed to be donors unless they explicitly opted out of donation; (iv) *mandated enrollment*: citizens did not have the choice of whether to become a donor and this decision was instead determined by a medical examiner. Participants were randomly assigned to one of the five possible comparison conditions in which at least one option involved a default: (1) mandated choice vs. nonenrollment, (2) mandated choice vs. automatic enrollment, (3) nonenrollment vs. automatic enrollment, (4) nonenrollment vs. mandated enrollment, or (5) automatic enrollment vs. mandated enrollment. The two policy descriptions were presented on a single page, with the order of the two policies counterbalanced across participants.

After reading the policies participants responded to two blocks of items, representing assessments of paternalism and justifiability. For the paternalism block participants indicated how "imposing" and "hands off" (reverse scored) the policy was for each state (e.g., "How imposing is State A's policy?"). For the justifiability block participants indicated the extent that the policies in State A and B were each "in need of a legitimate justification by policymakers," "in need of a convincing rationale," and should "require compelling reasons to be implemented." All items were made on 7-point scales (1 = *not at all*, 7 = *a great deal*), and were combined to form indices of paternalism ($\alpha = 0.72$) and justifiability ($\alpha = 0.84$). Each block of items was shown on a separate page, and the order of the two blocks was counterbalanced across participants. We also randomized the order of items within each block of questions.

At the end of the study, participants reported their age, gender, ethnicity, and political orientation. We also included a two-item general direction check to probe if participants had read our materials thoroughly (Oppenheimer et al., 2009). Seventy-four percent of participants answered both direction check questions correctly. For purposes of simplicity and completeness we include all participants, including those who failed the direction check; restricting the analysis to only participants who correctly answered both direction check items returns similar results (see Table 1 in the Supplemental Materials).

Results

We expected automatic enrollment to be rated as more paternalistic, and requiring greater justification, than nonenrollment. Furthermore, automatic and nonenrollment should be rated as falling between mandated enrollment (most paternalistic) and mandated choice (least paternalistic). To test these predictions we conducted paired *t*-tests for each of our five comparison conditions.

As displayed in Table 3, our results largely conform to this pattern. Corroborating our predictions, mandated choice and nonenrollment policies were viewed as the least paternalistic, followed by automatic enrollment and then mandated

---

[7]At the end of the study participants reported their familiarity with cornea transplant policies on a 7-point scale (1 = *very unfamiliar*, 7 = *very familiar*). As expected, most participants were unfamiliar with cornea transplant policy, with 86% of respondents reporting below the midpoint of the scale. Furthermore, rated familiarity with cornea transplant policies did not reliably moderate any of the results we report above (*p*-values ranged from 0.14 to 0.94).

Table 3: Paternalism and Justifiability Ratings in Study 2

	Mandated Choice (MC)	Non-Enrollment (NE)	Automatic Enrollment (AE)	Mandated Enrollment (ME)	Cohen's d
How Paternalistic?					
MC vs. NE	3.76 (1.59)	3.37 (1.85)			−0.13
MC vs. AE	2.86 (1.46)		5.31 (1.55)		0.91***
NE vs. AE		2.76 (1.62)	5.26 (1.62)		0.83***
NE vs. ME		2.76 (1.69)		5.71 (1.49)	1.02***
AE vs. ME			4.05 (1.22)	5.56 (1.65)	0.77***
Needs Justification?					
MC vs. NE	3.28 (1.67)	4.13 (1.83)			0.29†
MC vs. AE	2.42 (1.66)		5.34 (1.78)		1.11***
NE vs. AE		2.83 (1.88)	5.05 (1.99)		0.69***
NE vs. ME		2.81 (1.62)		5.93 (1.50)	1.17***
AE vs. ME			3.82 (1.84)	5.79 (1.57)	0.84***

Notes: Cells represent average responses in each condition, with standard deviations in parentheses. Cohen's d is calculated using the difference score between the two cells divided by the standard deviation in the difference score. Positive values for Cohen's d indicate results consistent with our hypotheses, whereas negative values indicate results inconsistent with our hypotheses. $^{\dagger}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

enrollment. All comparisons were significantly different from one another at $p < 0.001$, with the exception of the mandated choice vs. nonenrollment condition. Importantly, an automatic enrollment default was viewed as more paternalistic than a nonenrollment default, $t(41) = 5.37$, $p < 0.001$, $d = 0.83$.

A similar pattern was found for assessments of justifiability. Mandated choice and nonenrollment were lowest in justifiability, followed by automatic enrollment and mandated enrollment. All comparisons were significantly different from one another at $p < 0.001$, with the exception of the mandated choice vs. nonenrollment condition. Importantly, an automatic enrollment default was viewed as requiring greater justifications by the choice architect than a nonenrollment default, $t(41) = 4.46$, $p < 0.001$, $d = 0.69$.

Also as predicted, paternalism and justifiability scores were highly correlated. Across all observations we see a strong positive relationship between the two ratings (Pearson's $r = 0.66$, $p < 0.001$), with the correlation ranging from 0.51 to 0.73 across our comparison conditions (all p-values < 0.001). In short, a paternalistic policy is a policy in need of justification, and an automatic enrollment default is viewed as more paternalistic than a nonenrollment default.

Two other points are worth mentioning. First, we observe the largest differences in the two comparison conditions that are the furthest apart on our expected continuum (comparisons that are "two rungs" apart instead of "one rung" apart; i.e., rows MC vs. AE and NE vs. ME in Table 3). This finding suggests that participants were sensitive to gradations in paternalism and justifiability across the different comparison conditions in a manner consistent with our predictions. Second, and contrary to our expectations, participants gave roughly equal assessments to the mandated choice and nonenrollment default policies. One possible explanation for the equivalence between mandated choice and nonenrollment is that although mandated choice makes no presumptions on behalf of individuals, it does require individuals to make an explicitly choice — and foisting a choice upon individuals can be viewed as a small form of paternalism (Sunstein, 2014). Regardless, the lack of a meaningful difference between mandated choice and nonenrollment reinforces the more basic point we wish to make: even though nonenrollment and automatic enrollment defaults are formally equivalent choice regimes, automatic enrollment is viewed as considerably more paternalistic than nonenrollment (which in turn is viewed as no more paternalistic than a policy that requires individuals to actively state their preferences).

Study 3

In Study 3 we move away from a focus on default inferences and begin to explore the behavioral implications of the recommendation asymmetry. In particular, we examine how trust in the choice architect moderates default behavior for automatic and nonenrollment defaults. Although previous work has suggested that trust in the choice architect will likely play an important role in the effectiveness of behavioral policy design (Krijnen et al., 2017; McKenzie et al., 2018; Tannenbaum et al., 2017), these accounts do not specify what form this relationship will take. The recommendation asymmetry, however, makes a concrete prediction about the relative weight of trust on default behavior for different types of policy defaults. Since individuals tend to view automatic enrollment as more revealing about a choice architect recommends, feelings of trust in the choice architect should be more predictive of default acceptance for automatic enrollment than nonenrollment. To test this hypothesis, we presented participants with a hypothetical default decision in which we varied both the policy default as well as information about the choice architect as either trustworthy or untrustworthy. We pre-registered the design and analysis for Study 3 at `http://aspredicted.org/blind.php?x=qk47ir`.

Method

We recruited a sample of 1,001 U.S. participants from MTurk (55% female, average age = 37 years, range: 18–73 years) using the same recruiting and screening procedure as Study 1.

Our study used a 2 (default: non enrollment vs automatic enrollment) × 2 (trust: low vs high) between-participants design. Participants were asked to imagine recently moving to a new country and confronting a decision whether to have their child vaccinated for a particular disease. We focused on vaccination decisions because public trust plays an important role for such decisions (Larson et al., 2011, 2014). In the high-trust condition, participants read that this county was "known for its strong democratic institutions and low levels of corruption among government officials." In the low-trust condition, the country was described as having "recently been plagued with a string of political scandals involving bribes and other forms of corruption among elected government officials." Participants were then asked to imagine having a child who attends a public school in country A. Participants in the nonenrollment and automatic enrollment condition then read the following, respectively:

> *Nonenrollment default*:
>
> As part of its public health policy, the government of country A has recently enacted the following vaccination policy at all public schools: students only receive Vaccination X when a parent explicitly requests that their child receive the vaccine. (This vaccine was recently developed by a pharmaceutical company headquartered in country A)
>
> Imagine your child attends a public school in country A, and so your child is currently scheduled to not receive the vaccine.
>
> ☐ If this is fine, click here
>
> ☐ If you would prefer that your child receive the vaccine, click here

> *Auto-enrollment default*:
>
> As part of its public health policy, the government of country A has recently enacted the following vaccination policy at all public schools: all students receive Vaccination X unless a parent explicitly requests that their child not receive the vaccine. (This vaccine was recently developed by a pharmaceutical company headquartered in country A)
>
> Imagine your child attends a public school in country A, and so your child is currently scheduled to receive the vaccine.
>
> ☐ If this is fine, click here
>
> ☐ If you would prefer that your child not receive the vaccine, click here

After reading this information, participants then made a decision to either stay with the default option or to instead have their child vaccinated/not vaccinated (depending on condition). As a manipulation check, on the next page participants were re-shown the description of country A and asked to rate on 7-point scales (1 = *not at all*, 7 = *very much so*) the degree that "government officials in country A probably make policies with their citizens' best interests in mind," "government officials in country A probably make policies based on what is in their own personal interests, rather than what is good for citizens" (reverse-scored), "government officials in country A appear trustworthy," and "I am suspicious of the motives of government officials in country A" (reverse scored). We averaged these items to form a single measure of trust ratings (Cronbach's $\alpha = 0.92$), with higher numbers indicating greater trust. At the end of the study we asked participants to report their age and gender.

Results

First, as a manipulation check, we observe that participants viewed government officials as more trustworthy in the high-trust condition ($M = 5.07$, $SD = 1.23$) than in the low-trust condition ($M = 2.57$, $SD = 1.31$), $t(999) = 31.10$, $p < 0.001$, $d = 1.97$. We also observe standard default behavior, with 79% of participants choosing vaccination for their child under an automatic enrollment default compared to 68% under nonenrollment, $Z = 3.98$, $p < 0.001$.

Our main prediction is that trust in the choice architect would have a greater impact on default decisions under automatic enrollment than nonenrollment. To test this we conducted a logistic regression with default acceptance as our outcome variable (0 = leave default option, 1 = stay with default option), and our two conditions and their interaction term as predictor variables (trust condition: 0 = low trust, 1 = high trust; default condition: 0 = nonenrollment, 1 = automatic enrollment). As predicted, we observe a positive interaction effect, $b = 0.63$, SE = 0.29, $p = 0.032$. Under nonenrollment, default behavior did not reliably differ between the low-trust condition (34% stay with default option) and the high-trust condition (29% staying with default option), $Z = 1.17$, $p = 0.243$. Under automatic enrollment, participants were more likely to display default behavior in the high-trust condition (83%) than in the low-trust condition (76%), though this difference was only marginally significant, $Z = 1.83$, $p = 0.067$.

Follow-up Study

The results of Study 3 provide initial support for a key prediction made by the recommendation asymmetry hypothesis. Since nonenrollment defaults are viewed as relatively uninformative about what a choice architect recommends, such inferences are less likely than automatic enrollment defaults to be qualified by the choice architect's degree of trustworthiness. Although the results of Study 3 provide causal evidence for this prediction, one limitation is the stylized nature of the decision task.

In the Supplemental Materials we report the results from two additional studies (Studies S4 and S5) that examined whether the results of Study 3 generalize to other decision settings. In Study S4 we presented a new sample of MTukers ($N = 750$) with one of three hypothetical decisions: (i) whether to become an organ donor upon their death; (ii) whether to enroll into a credit card payment plan that withdrew the minimum amount due each month from their bank account; or (iii) whether to join a company's savings plan in which 5% of monthly earnings were set aside for retirement. Participants were assigned to either an automatic enrollment or nonenrollment default, and made a decision to either stay with the default option or change to the alternative course of action. Afterwards we asked them to rate their general degree of trust in the choice architect.

Figure 1 shows the main results, in which the predicted probability of staying with the default option is plotted as a function of trust in the choice architect under automatic and nonenrollment defaults. Consistent with the results of Study 3, we observe that in all three settings the marginal effect of trust on choice was higher under automatic enrollment (as

14

Figure 1: Study S4 Results
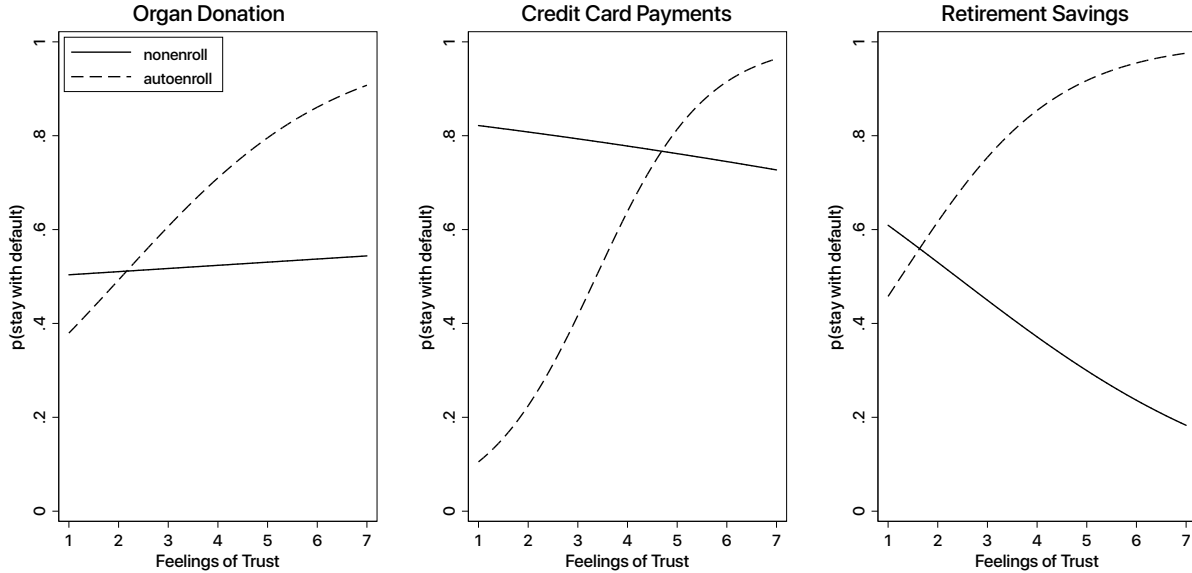
Notes: Lines represent predicted probabilities calculated from average marginal effects from the logit models specified in Study S4.

shown by the positive slope of the lines in Figure 1) than nonenrollment. In all three settings the interaction effect in the predicted direction, and across settings is significant at $p < 0.001$.

In Study S5, we presented a new sample of 200 MTurkers with three hypothetical decisions: (i) the same vaccination decision used in Study 3, (ii) a bank customer privacy policy on sharing personal and financial information with its affiliate agencies, and (iii) a retailer policy on extended warranties for customer product purchases. Similar to Study 3, we also randomly varied information about the institution or organization to suggest that the choice architect was trustworthy or untrustworthy, as well as whether the policy default was one of automatic enrollment or nonenrollment. In all three settings the interaction effect in the predicted direction, and across settings is significant at $p < 0.001$. In summary, we find that a key prediction made by our recommendation asymmetry hypothesis — that trust in the choice architect is more positively associated with default behavior under automatic enrollment than nonenrollment — generalizes to other hypothetical decision settings.

Study 4

Study 4 examines the link between default inferences, trust, and behavior in a field experiment with real stakes. The study took place in a classroom setting, where university students decided how to schedule three research reports over the course of the quarter. The three reports constituted a significant portion of the student's overall grade in the course (25%), so any decision impacting performance would be consequential. All students had the option of imposing "costly" deadlines by committing to a binding deadline schedule to submit each report before the end of the course. The default deadline schedule varied across students and served as our key manipulation.

Our study takes its design from a previous experiment by Ariely and Wertenbroch (2002). In their experiment, students either did or did not have the option to impose costly deadlines on themselves. When students were provided with the opportunity to impose deadlines, they usually did so and their grades improved as a result. Compared to those

in the control group — where all three reports were due by the last day of class, but could always be turned in early — students who imposed external deadlines on themselves were less likely to procrastinate, and as a result wrote better papers. What makes this finding interesting is that the policy in the control group (i.e., setting all deadlines to the final day of class) is in some sense the optimal allocation strategy. Ariely and Wertenbroch put it the following way: "By setting their deadlines as late as possible, the students would have the most time to work on the papers, the highest flexibility in arranging their workload, and the opportunity to learn the most about the topic before submitting their papers" (p. 220). However, students often procrastinate and so committing to a stricter deadline schedule ends up improving performance.

The current study adapts this design by providing all students with the option of imposing costly deadlines, and instead manipulates the default deadline schedule (Ariely and Wertenbroch's participants were either given no choice at all or required to choose without a default setting). For half of the students the default was an even allocation schedule, with the deadlines "staggered" throughout the course. For the other half, the default was for all deadlines to be "lumped" together at the end of the course.[8] Thus, our study represents what Harrison and List (2004) call a "natural field experiment" as students did not have knowledge of being involved in an experiment.

Lastly, we note that our two default conditions do not perfectly map onto our distinction between automatic enrollment and nonenrollment. Recall that an automatic enrollment default is one that alters an individual's state of affairs in the absence of choice. In Study 4 both the staggered and lumped defaults alter the current state of affairs, since students must turn in their papers at some point during the course. However, based on our results from Study 2 we reasoned that the basic asymmetry should hold given that the staggered default likely plays a particularly active role in structuring students' deadlines, and thus the same mechanism governing the recommendation asymmetry should still operate in this context. To verify this assumption, we conducted a pilot study[9] asking participants to evaluate the two class policies. The results clearly indicated that students viewed the staggered default as more interventionist than the lumped default. Therefore, we expected that students would be more likely to infer an instructor's reasons and beliefs from the staggered default than from the lumped default. We further predicted that drawing inferences from the default deadline schedule would be associated with greater default behavior, and that general feelings of trust in the choice architect should be more predictive of default decisions for the staggered default than for the lumped default.

Method

Our sample consisted of 169 undergraduate students enrolled in an upper-division course on judgment and decision making.[10] Most students were psychology majors, and the rest majored in disciplines primarily within the social

---

[8]We note that the design of Study 4 does not allow us to directly establish a recommendation asymmetry between the staggered deadlines and lumped deadlines conditions, since we do not include a no-default control condition. As discussed in Study 1, to clearly establish a recommendation asymmetry one needs to account for prior beliefs about a choice architect's preferences (which can be achieved by comparing responses to a control condition). Given that our sample size was fixed in Study 4 (i.e., based on the number of students enrolled in the course), including a control condition would substantially reduce statistical power to detect differences across conditions. Furthermore, recall that in Study 1 the modal response in the control condition was always that the policy "tells me nothing" about a choice architect's preferences. Given that most participants did not draw strong inferences about the choice architect in more familiar contexts such as organ donation and retirement savings — with stronger social norms about what most people prefer — we thought it unlikely that participants would have strong prior beliefs about an instructor's preferences over paper deadlines. We thus settled for a weaker test of the recommendation asymmetry (though with greater statistical power) by comparing inferences from the staggered deadlines default directly to those from the lumped deadlines default, rather than against a no-default control condition.

[9]We recruited a sample of 169 students (drawn from the same student population used in Study 4) to imagine enrolling in a university course and were told that the course required them to write three research papers. Half of students were provided with a description of the staggered default policy, and the other half were provided with a description of the lumped default policy. Students then rated on 7-point scales (1 = *strongly disagree*, 7 = *strongly agree*) whether the policy was "highly structured for students," "very 'hands on' regarding the deadlines," "highly unstructured for students," and "very 'hands off' regarding the deadlines." The latter two items were reverse coded and then all items were averaged into a single index ($\alpha = 0.72$). As expected students viewed the staggered default as more structured ($M = 5.03$, $SD = 0.99$) than the lumped default ($M = 3.10$, $SD = 1.38$), $t(167) = 10.43$, $p < 0.001$, $d = 1.60$.

[10]The first author was the instructor for the course. The study used a double-blind design: students were randomly assigned to deadline conditions upon completing a web survey, and their papers were evaluated by course teaching assistants who were blind to conditions and to the study hypothesis.

Figure 2: Study 4 Timeline

Staggered Default
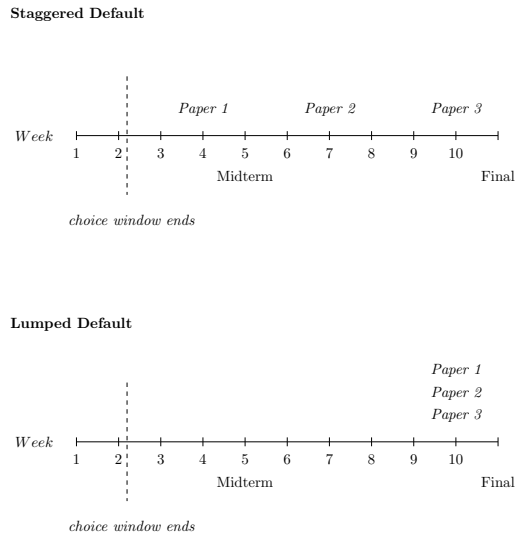
Lumped Default

Notes: The figure depicts the default schedule for staggered and lumped conditions. The choice window ended at the end of week 2. In the staggered default, the deadlines for reports 1, 2, and 3 were set for weeks 4, 7, and 10. In the lumped default, the deadlines for reports 1, 2, and 3 were set for week 10. For any deadline, students had up until the last day of the week (Sunday at midnight) to electronically submit their report. The midterm exam took place at week 5, and the final exam took place a week after the last day of class.

sciences. There were 80 seniors, 72 juniors, and 17 sophomores enrolled in the course.

On the first day of class the instructor went over the course syllabus and all course requirements. One requirement involved writing three research reports, each 2–3 pages in length, that comprised 25% of the student's total grade in the course. Students were told that since the subject of the course was decision making, they would have the opportunity to "make some real decisions about how [they] would like the course to be structured" and had the opportunity to arrange any set of due dates for submitting their research reports. Similar in design to Ariely and Wertenbroch (2002), the following constraints were imposed: (i) all papers had to be submitted by the last day of the course, but could always be submitted earlier, (ii) students had to indicate their deadline preferences by the end of the second week, (iii) once the 2-week window closed deadline schedules could no longer be changed or revised, and (iv) deadlines were binding, with late papers penalized roughly 1% of the student's overall course grade per day late. Students were explicitly told that they would not receive feedback or grades on the reports before the end of the course. Lastly, the assignment prompt for each paper was available online so that students had complete information about the task when scheduling their deadlines.

Following the first lecture students were provided with a url link directing them to an online web form for establishing their deadline preferences, and were randomly assigned to one of two conditions (see Figure 2 for an overview). Roughly half of students were assigned to the *staggered deadlines* condition, where the default was to submit report 1 at the end of week 4, report 2 at the end of week 7, and report 3 at the end of week 10. The other half of students were assigned to the *lumped deadlines* condition, where the default was for all three reports to be due on the final week of class (week 10). When completing the online form, students first indicated whether they wished to stay with the default or create their own custom schedule, and those who chose the latter were directed to a second webpage to indicate their specific preferences. Other than the default setting, the information provided in both conditions was identical.

After students indicated their preferences online, they were then asked to complete a series of questions ostensibly

Table 4: Study 4 Questions.

Decision Difficulty
    1. I found this decision easy. (reverse scored)
    2. I found this decision difficult.
    3. I found this decision to be confusing.
    4. I felt confident in my decision. (reverse scored)

Default Inferences
    1. The instructor probably thinks that students should be enrolled in the default deadline schedule.
    2. The instructor probably thinks that most students would achieve a better grade by being in the default deadline schedule.
    3. The instructor probably thinks that most students themselves would prefer to be in the default deadline schedule

Trust in Instructors
    1. In general, I believe that my instructors typically have my best interests in mind
    2. In general, I believe that my instructors typically are out to punish their students (reverse scored)
    3. When it comes to course policies, I believe that my instructors typically have a solid understanding of how to best structure the course
    4. When it comes to course policies, I believe that my instructors typically have no idea what they are doing (reverse scored)

Notes: Participants rated their agreement with each statement on 7-point scales (1 = (*Strongly Disagree*, 7 = *Strongly Agree*).

designed to "provide the instructor with feedback on how to best structure courses in the future." Students were then presented with items measuring decision difficulty, default inferences, and general trust in their course instructors, in that order. For default inferences, the first two items assessed inferences of implicit advice and the last item assessed inferences of majority preference. All items are displayed in Table 4. Students indicated their agreement with each statement on a 7-point scale (1 = *Strongly Disagree*, 7 = *Strongly Agree*).

Later on, when students completed their research reports, they submitted their papers through a course website which provided an electronic time-stamp of the submission. At the end of the course, two teaching assistants (blind to the research hypothesis) each graded half of the reports. Also, prior to the final week of class, students were asked to complete an online survey in which they provided evaluations and feedback on the course. Of relevance here were a set of evaluations regarding the deadline scheduling policy. One item assessed evaluations about the policy ("Regarding the research report deadlines, I liked having the option to schedule my deadlines as I wished") and one item probed for suspicion ("Any other comments, or anything you found curious, about the deadline scheduling?"). Participants responded to the first item on a 5-point scale (1 = *Strongly disagree*, 5 = *Strongly Agree*) and the second item was open-ended.

Results

Of the 169 students enrolled in the course, 21 never completed the online survey necessary to register their deadline preferences, and were assigned to one of the two default schedules by the instructor. We were unable to include these participants in the analysis as they never confronted the primary decision to stay with the default schedule, nor did they provide responses to our default inference items. An additional nine students completed the deadline scheduling task twice;[11] for the analysis we used these students' last decision, as the course policy was to tie deadline schedules to the final decision registered by the student. Finally, in the end-of-semester survey four participants reported suspicion about the deadline scheduling policy being an experiment. We excluded these participants from the analysis, yielding a final sample of 144 students; including suspicious participants does not change the direction or significance levels for any

---

[11]Of the nine participants who completed the deadline scheduling task twice, seven participants were exposed to both default conditions. For completeness and simplicity we include these participants in the analyses; excluding them from the analysis does not change the direction or significance levels of any results we report.

Figure 3: Deadline Preferences by Default Condition (Study 4)
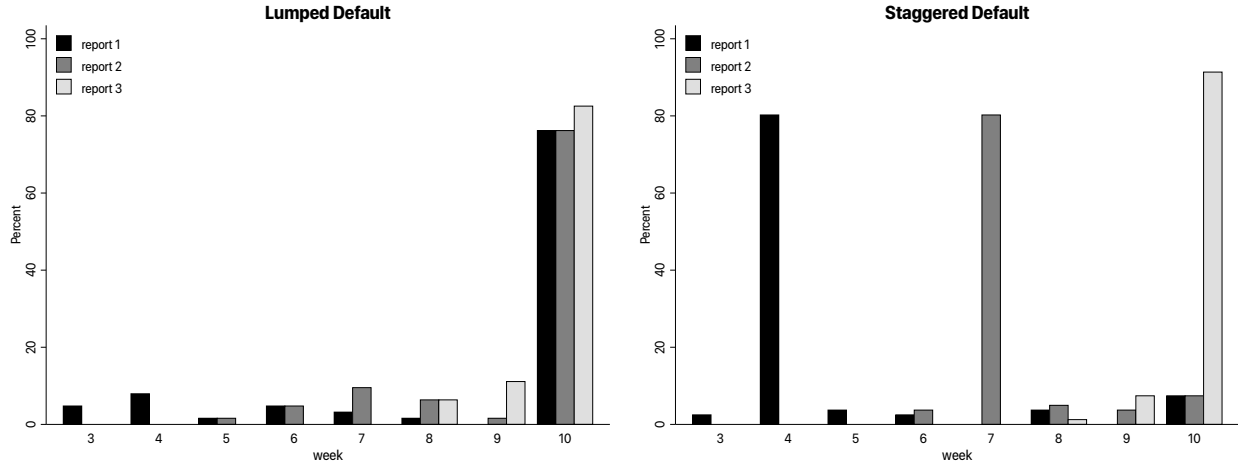
Notes: Deadline preferences under the lumped deadlines default (left panel) and staggered deadlines default (right panel). The x-axis is the course week that the students chose, for each report, to impose a deadline. The y-axis is the percent of participants, separately for each report, who selected a particular deadline.

results we report below.

Evaluations of the Decision Task   Although not our primary outcome of interest, it is instructive to examine students' evaluations of the decision context. Immediately after registering their default decision students were asked to rate the difficulty of the task on four items, which we averaged and combined into a single index ($\alpha = 0.86$). The majority of students did not find the task difficult, with 81% responding at or below the midpoint. We also asked students at the end of the course to rate how they liked having the option to schedule their own deadlines (on a 5-point scale from *strongly disagree* to *strongly agree*), and the clear majority of students (90%) were either neutral or positive towards the policy (responding at or above the midpoint). Overall students did not appear to find the task difficult or confusing, and were largely positive about the opportunity to set their own paper deadlines.

Default Decisions   Given that most students enjoyed having the ability to choose their own deadlines and viewed the task as not particularly difficult, one might expect few students to stay with the default deadline schedule. However, we observed that that a large majority of students stayed with the default option (77% of students in the staggered deadlines condition and 73% of students in the lumped deadlines condition). As a result, the distribution of deadline allocations differed substantially across the two default conditions. Shown in Figure 3, the distribution for the staggered default shows large spikes at weeks 4, 7, and 10, corresponding to the default schedule. For the lumped default, the mass of the distribution is at week 10, again corresponding to the default schedule. Accordingly, students in the lumped deadline had later deadlines than students in the staggered default for reports 1 and 2 (both *p*-values < 0.001 by a Mann-Whitney *U* test). Only for report 3, when both defaults assigned a deadline of week 10, was there no statistically significant difference in deadline allocation[12] ($p = 0.101$ by a Mann-Whitney *U* test).

---

[12]If we compare the two distributions using a Kolmogorov-Smirnov test, an alternative non-parametric test to the Mann-Whitney, we observe a similar pattern. The distributions for reports 1 and 2 are reliably different from one another across default conditions (both *p*-values < 0.001), but the distributions for report 3 are not reliably different from another ($p = 0.922$).

Default Inferences   We predicted that students would be more likely to view the staggered default than the lumped default as informative about the instructor's reasons and motivations. Our inference measures involved two items measuring whether the default signaled implicit advice by the instructor and one item measuring whether the default signaled majority preference (what most students would likely choose for themselves if asked). Unlike our previous studies which measured default inferences, in this context the majority preference item was weakly correlated with the implicit advice items ($r = 0.11$ and .06, respectively). For this reason we chose to average the two implicit advice items into a single index (Spearman-Brown inter-item reliability $r = 0.69$) and analyze the majority preference item separately.

As expected, students were more likely to infer implicit advice from the staggered default ($M = 4.34$, $SD = 1.37$) than from the lumped default ($M = 3.44$, $SD = 1.26$), $t(142) = 4.03$, $p < 0.001$, $d = 0.68$. Interestingly, the opposite pattern emerged for majority preference. Students were marginally less likely to infer majority preference from the staggered deadline default ($M = 4.37$, $SD = 1.74$) than from the lumped deadline default ($M = 4.87$, $SD = 1.32$), $t(141) = 1.89$, $p = 0.061$, $d = 0.32$. Although we did not predict this latter finding, it is perhaps not surprising that students would assume most of their peers prefer turning in their papers later rather than sooner.
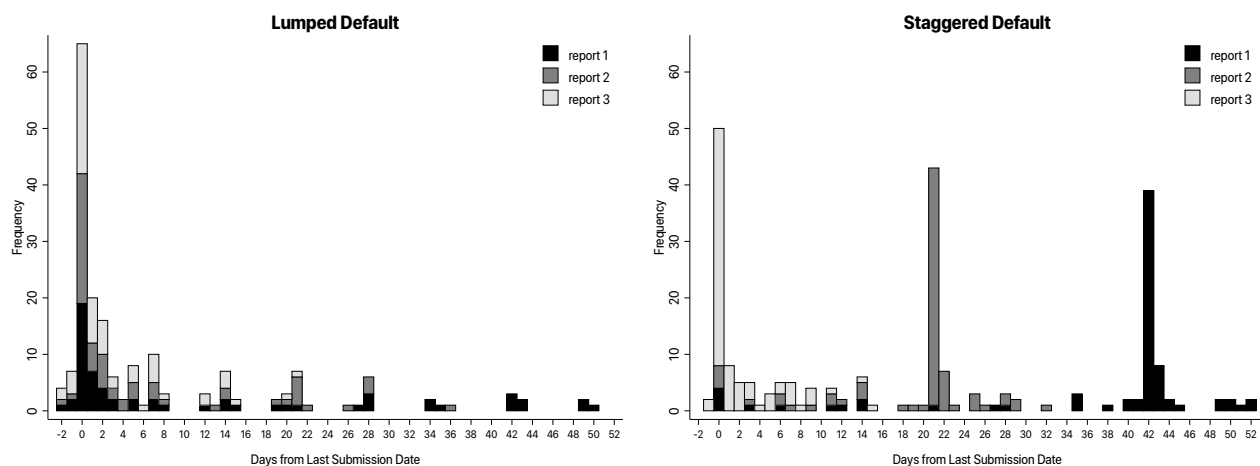
Crucially, we also observed that inferences of implicit advice were associated with default behavior. To test this we conducted a logistic regression using default behavior as our outcome variable (0 = leave default option, 1 = stay with default option) and ratings of implicit advice and majority preference as predictor variables (for all analyses, we implement robust standard errors). When the model included only implicit advice as a predictor, we find that students who made stronger inferences of implicit advice were also more likely to stay with the default option, $b = 0.40$, SE = 0.16, $p = 0.010$. When the model included only majority preference as a predictor, we observed a similar but weaker pattern, $b = 0.23$, SE = 0.13, $p = 0.072$. When both implicit advice and majority preference were included in the model, only implicit advice was a reliable predictor, $b = 0.38$, SE = 0.16, $p = 0.017$. Inspection of the average marginal effects indicate that a one-unit increase in ratings of implicit advice (on our 7-point scale) was associated with a roughly 7 percentage-point increase in default behavior.

Trust in Choice Architect and Default Behavior   We predicted that feelings of trust would be a stronger predictor of default acceptance for the staggered deadline default than the lumped deadline default (since the staggered default was more likely to be viewed as an implicit recommendation by the instructor). We first combined and averaged our four trust items to create an index of general trust in instructors ($\alpha = 0.74$). We also verified that ratings of trust did not reliably differ between the staggered default condition ($M = 5.93$, SD = 0.92) and the lumped default condition ($M = 5.89$, SD = 0.83), $t(141) = 0.26$, $p = 0.792$, $d = 0.04$.

We next conducted a logistic regression using default acceptance as our outcome variable, and default condition (0 = lumped deadlines, 1 = staggered deadlines), trust ratings, and the interaction term as our predictor variables (again implementing robust standard errors). We observed a reliable interaction between trust and policy defaults on decisions to stay with the default option, $b = 1.04$, SE = 0.46, $p = 0.023$. Consistent with the results from our earlier studies, examining the average marginal effects for each condition we find that trust was a stronger predictor of default acceptance for staggered deadlines (12.1% marginal effect, $p = 0.007$) than for lumped deadlines (-6.3% marginal effect, $p = 0.359$).

Consequences Tied to Default Decisions   Lastly, we examined whether decisions to stay with or leave the default had a real impact on course outcomes. We used course grades and paper submission times (a proxy for procrastination) to address this question. First, the default option did not appear to impact student grades in any meaningful sense. We failed to observe a reliable difference in course grades between students in the staggered deadlines default ($M = 73.01$,

Figure 4: Distribution of Paper Submissions by Default Condition (Study 4).

Lumped Default

Staggered Default

report 1
report 2
report 3

Frequency

Days from Last Submission Date

Notes: Paper submission times for each report, by default condition. The x-axis represents submission times (in days) from the final day of the course. Negative scores represent papers submitted after the final deadline. The y-axis is the number of submissions per report for a given day.

$SD = 14.27$) and lumped deadlines default ($M = 75.41$, $SD = 11.67$), $t(142) = 1.08, p = 0.282, d = 0.18$. When examining only grades for the research reports we do find slightly lower scores in the staggered default ($M = 75.00$, $SD = 21.29$) than in the lumped default ($M = 80.59$, $SD = 15.82$), but this difference was small and only marginally significant, $t(142) = 1.74, p = 0.083, d = 0.29$.

One reason why the default did not meaningfully impact course grades may have been because students did not in fact procrastinate. Students in the lumped default may have allocated equal amounts of time to each paper without the "help" of an external deadline, in which case the default would truly be inconsequential. The distribution of submission times, however, suggest that this was emphatically not the case. Shown in Figure 4, there were large spikes in submission times that correspond to the deadlines for each default condition — students submitted their research reports close to the deadline, whatever that deadline happened to be. To test for differences across the two defaults, each submission time was scored by taking its distance in minutes from the absolute deadline for all reports (midnight on the end of week 10). For report 1, students in the lumped default turned in their reports on average 25.42 days later than students in the staggered default, and 12.33 days later for report 2 (both p-values < 0.001 by a Mann-Whitney U-test). Only for report 3, when both defaults set the due date for the last day of class, was there no reliable difference in turn-in times ($p = 0.586$ by a Mann-Whitney U-test).

Discussion

Study 4 yielded a number of findings. First, the large majority of students stayed with the default option, regardless of whether the default was to stagger deadlines throughout the course or lump them altogether on the final day of the class. This large default effect also had downstream consequences on students' behavior, with most students submitting their reports close to the deadline (whatever that deadline happened to be).

Second, students were more likely to view the staggered deadlines default as an implicit recommendation by the instructor than the lumped deadlines default. Importantly, inferences of implicit advice predicted whether students chose to stay or leave the default option. The more that students viewed the default as an endorsement by the instructor, the more likely they were to stay with that option.

Third, trust in the instructor moderated default behavior. Feelings of trust was a stronger predictor of default behavior when the default was to stagger deadlines rather than lump all deadlines together at the end of the course. This pattern matches that observed in Studies 3, S4, and S5, and is consistent with a recommendation asymmetry in favor of automatic enrollment defaults.

Overall, the results of Study 4 suggests that default inferences can play an important role in determining default behavior. Study 4 is also noteworthy because many of the elements that should, from a rational actor standpoint, eliminate default effects were in fact present. Direct transaction costs were minimal, since students could easily leave the default by completing a short survey at their convenience. Students were experienced in the task domain and surely have made time allocation decisions for papers from previous courses (all students in the class were at least Sophomores). Students were presumably motivated to decide well, and did not report finding the task overly burdensome or difficult. All of these elements have been proposed as factors that should reduce decision biases (Shafir and LeBoeuf, 2002), yet we still observed a large default effect. However, such an analysis fails to account for possible informational differences between default options (Sher and McKenzie, 2008, 2006). To the extent that defaults differ in pay-off relevant information (such as the recommendation of a potentially-knowledgeable choice architect), decision makers should update their beliefs based on the default. Furthermore, these updated beliefs can also lead to normatively-justified differences in preferences across default options. The results of Study 4 suggest that students draw different inferences about an instructor's reasons and beliefs from course policy defaults, and these inferences shape their deadline preferences.

General Discussion

Default options can convey information relevant to making a default decision. In this paper we document that automatic enrollment defaults are often thought to convey information about what course of action is sensible or best, whereas nonenrollment defaults are relatively uninformative about what a choice architect recommends. This recommendation asymmetry arises because individuals assume automatic enrollment reflects a decision by the choice architect who selected the default for goal-directed reasons, whereas nonenrollment often reflects simple inaction without any supporting reasons. Furthermore, we test and corroborate a key implication of the recommendation asymmetry, namely that feelings of trust in the choice architect are more predictive of default acceptance for automatic enrollment than nonenrollment defaults.

Naturally, our studies have several limitations. First, our studies draw on participant populations from Mechanical Turk (Studies 1–3) and undergraduate university students (Study 4). It is unclear whether other populations make default inferences at similar rates, or in a similar manner, to those studied here. However, field experiments on the effects of defaults for retirement savings plans have used large companies with diverse workforces, and they find behavioral patterns that are consistent with the recommendation asymmetry we document here (e.g., Madrian and Shea, 2001). Furthermore, Brown et al. (2011) directly surveyed employees about their default retirement savings option, and many reported viewing the default as a recommendation by their employer. A fruitful avenue for future research would be to examine how specific populations and situational features interact to lead individuals to draw inferences from default options (see Krijnen et al., 2017, for one possible approach). Second, in our studies participants made inferences about default options in response to questions posed by an experimenter, so it is unclear to what degree participants generate such inferences spontaneously. Although this is a reasonable concern, previous laboratory studies using more subtle procedures have found that participants spontaneously infer different information from logically equivalent frames (Sher and McKenzie, 2006). Furthermore, we also note that in our studies which directly examined default behavior we either did not probe for default inferences at all (Studies 3, S4, and S5), or asked participants to report their inferences

only after making a default decision (Study 4). Therefore, the effects we observe on behavior are unlikely to be an artifact of prodding respondents to report their inferences about a default option.

An additional concern involves the issue of stimulus selection (Clark, 1973; Wells and Windschitl, 1999). Naturally, the inferences that participants draw from a default option will vary from context to context, and so it is reasonable to ask whether the recommendation asymmetry documented here is an artifact of the particular settings we examined. Stimulus selection is always a concern in psychological research, and we have attempted to concern this issue by testing for a recommendation asymmetry across a number of contexts that involve automatic enrollment and nonenrollment defaults. Across studies this list includes organ donation, cornea donations, retirement savings, bank overdraft protection, medical testing for an infection, vaccination decisions at public schools, credit card repayment plans, customer privacy and information sharing decisions, customer warranty purchasing decisions, and students selecting deadlines for their term papers. Although these contexts differ along a number of dimensions (such as whether the choice architect is a government or for-profit entity), across settings we find evidence consistent with a basic recommendation asymmetry between automatic and nonenrollment defaults.

Our findings contribute to a growing body of research showing how policies influence behavior through the social messages they communicate (Beshears et al., 2017; Everett et al., 2015; Gneezy and Rustichini, 2000; McAdams and Nadler, 2005; McKenzie et al., 2006; Tannenbaum et al., 2013). More generally, our studies suggest that policy reactions can be thought of as a tacit social interaction taking place between a choice architect and the targets of the choice architecture (Grice, 1975; Krijnen et al., 2017). That is, individuals appear to interpret and explain choice architecture policies by relying on the same social cognitive tools that they use to explain and understand everyday social behavior (Malle, 2004). That default effects are partly social-psychological in nature has a number of interesting policy, managerial, and organizational implications. The following sections represent a partial list.

Implications for Libertarian Paternalism

Our account complicates the arguments for libertarian paternalism. Libertarian paternalists argue that managers and choice architects should create policies that nudge behavior in welfare-enhancing directions as long as they do not meaningfully obstruct individual choice (Camerer et al., 2003; Thaler and Sunstein, 2003). Since defaults have a large influence on behavior, and do not alter the choice set, they are often presented as prime examples of sound libertarian paternalistic policies (Thaler and Sunstein, 2008). Advocates of libertarian paternalism defend their prescriptions by appealing to two basic premises. First, all policies impact choice; this is an inescapable consequence of the fact that preferences are highly constructed and sensitive to context. Second, since all policies impact choice, policy should influence behavior in ways that enhance welfare (Thaler and Sunstein, 2003).

The current findings do not refute either of these claims, but do qualify them. On the first point, the degree to which defaults impact choice will depend partly on the information they provide. Since automatic enrollment defaults send especially strong signals of implicit advice, they actively influence default decisions in a way that nonenrollment defaults usually do not (or at least not to the same degree). For similar reasons, however, automatic enrollment is also potentially less paternalistic than nonenrollment. Automatic enrollment may play a more active role in shaping default decisions, but it is a participatory role — a suggestion that individuals either accept or reject. To the extent that nonenrollment defaults influence behavior through means other than rational persuasion (e.g., through inattention), using them to shape behavior is arguably more paternalistic (Hausman and Welch, 2010).

Another complication is that whenever defaults such as automatic enrollment convey information about what people ought to choose or what most people do choose, they have the potential to stigmatize non-default behavior (Glaeser, 2006). For example, individuals can be reluctant to opt-out of routine testing for sexual transmitted infections because doing so implies that one has "something to hide" (Young et al., 2009). That defaults can stigmatize behavior

would certainly place a meaningful burden on those who wish to leave the default, and would be contrary to the tenets of libertarian paternalism. Our account can help managers and choice architects understand when defaults may inadvertently stigmatize non-default choices.

Potential Downsides of Automatic Enrollment

Actively setting defaults can sometimes do more harm than good. Consider, for example, that automatic enrollment into a retirement savings plan substantially increases the number of employees who save for retirement. While virtually everyone would agree this is a positive outcome, the default allocation may often be too conservative. Recent surveys of companies using automatic enrollment savings plans find that many companies arguably set savings rates that are too low and use investment vehicles that are too conservative (Choi et al., 2004). In cases where automatic enrollment defaults are especially strong, this means that some employees who stay with the default might have eventually chosen a more aggressive savings plan on their own. Ironically, over the long run employees may accumulate *less* wealth under an automatic enrollment default when the policy is not properly calibrated to the needs of employees (for a discussion, see Choi et al., 2004).

More generally, there is also the possibility that defaults can undermine intrinsic motivation and increase feelings of regret (Botti and McGill, 2006). Passive acceptance of a default option leads to weaker identification with the accepted course of action (Cioffi and Garner, 1996) and inaction often leads to increased feelings of regret over the long run (Gilovich and Medvec, 1995). Indeed, university employees in Illinois who stayed with the default pension plan were more likely to report regret later on than employees who actively chose a different plan (Brown et al., 2011). This was the case even when controlling for outcomes: employees who failed to indicate a preference, and were therefore placed into the default savings plan, were more likely to report regret than employees who explicitly indicated a preference for the default option (35% vs. 14%). However, all employees in this study had the same default savings plan, so it is unclear if feelings of regret are associated with specific default options or are simply a result of passive default behavior. Nonetheless, choice architects may wish to consider how the strategic selection of a default option may have spillover effects on an individual's subsequent motivation and decision making.

Both of these issues highlight the complexities in setting defaults in ways that enhance welfare. One recommendation commonly put forth is to educate individuals about the decision task, such as increasing financial literacy for retirement savings. Default behavior has been tied to a lack of knowledge about the decision task (Agnew and Szykman, 2005; Agnew et al., 2008), so educational outreach programs should help to eliminate some of these problems (but see Hadar et al., 2011, for limitations to this approach). Interestingly, one study also found a positive correlation between inferring implicit advice from the default option and feeling incompetent in the task domain (Brown et al., 2011), suggesting that education programs may also reduce reliance of any implicit suggestions provided by the default. Other approaches include simplifying the decision task (Beshears et al., 2010b), mandating choice instead of having a default (Carroll et al., 2009), and creating defaults that adapt and learn from each individual's preferences ("smart defaults"; Goldstein et al., 2008). Managers and choice architects can supplement these approaches by identifying the preconditions that determine when defaults will be optimal or suboptimal (Beshears et al., 2010a; Carroll et al., 2009).

References

Abadie, A. and Gay, S. (2006). The impact of presumed consent legislation on cadaveric organ donation: a cross-country study. *Journal of Health Economics*, 25(4):599–620.

Agnew, J. R. and Szykman, L. (2005). Asset allocation and information overload: The influence of information display, asset choice, and investor experience. *Journal of Behavioral Finance*, 6(2):57–70.

Agnew, J. R., Szykman, L., Utkus, S., and Young, J. A. (2008). Literacy, Trust and 401 (K) Savings Behavior. Unpublished manuscript, College of William and Mary.

Ariely, D. and Wertenbroch, K. (2002). Procrastination, deadlines, and performance: Self-control by precommitment. *Psychological Science*, 13(3):219–224.

Baker, C. L., Saxe, R., and Tenenbaum, J. B. (2009). Action understanding as inverse planning. *Cognition*, 113(3):329–349.

Baron, J. and Ritov, I. (1994). Reference points and omission bias. *Organization Behavior and Human Decision Processes*, 59(3):475–498.

Bellman, S., Johnson, E., and Lohse, G. (2001). To opt-in or opt-out? It depends on the question. *Communications of the ACM*, 44(2):25–27.

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

Beshears, J., Choi, J. J., Laibson, D., and Madrian, B. C. (2010a). The limitations of defaults. Unpublished manuscript, Stanford University.

Beshears, J., Choi, J. J., Laibson, D., and Madrian, B. C. (2010b). Simplification and saving. Unpublished manuscript, Stanford University.

Beshears, J., Choi, J. J., Laibson, D., Madrian, B. C., and Weller, B. (2010c). Public policy and saving for retirement: The "autosave" features of the pension protection act of 2006. In Siegfried, J. J., editor, *Better Living Through Economics*, pages 274–290. Harvard University Press, Cambridge.

Beshears, J., Dai, H., Milkman, K., and Benartzi, S. (2017). Framing the future: The risks of pre-commitment nudges and potential of fresh start messaging. Technical report, Working paper.

Botti, S. and McGill, A. L. (2006). When choosing is not deciding: The effect of perceived responsibility on satisfaction. *Journal of Consumer Research*, 33(2):211–219.

Brown, J. R., Farrell, A. M., and Weisbenner, S. J. (2011). The downside of defaults. Unpublished manuscript, University of Illinois at Urbana-Champaign.

Camerer, C., Issacharoff, S., Loewenstein, G., O'Donoghue, T., and 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. I., Madrian, B., and Metrick, A. (2009). Optimal defaults and active decisions. *Quarterly Journal of Economics*, 124(4):1639–1674.

Chapman, G. B., Li, M., Colby, H., and Yoon, H. (2010). Opting-in vs opting-out of influenza vaccination. *Journal of the American Medical Association*, 304(1):43–44.

Childress, J. F. and Liverman, C. T. (2006). *Organ donation: Opportunities for action*. National Academy Press, Washington, D.C.

Choi, J. J., Laibson, D., Madrian, B. C., and Metrick, A. (2002). Defined contribution pensions: Plan rules, participant choices, and the path of least resistance. *Tax policy and the economy*, 16:67–113.

Choi, J. J., Laibson, D., Madrian, B. C., and Metrick, A. (2004). For better or for worse: Default effects and 401(k) savings behavior. In Wise, D. A., editor, *Perspectives on the economics of aging*, pages 81–121. University of Chicago Press, Chicago, IL.

Cioffi, D. and Garner, R. (1996). On doing the decision: Effects of active versus passive choice on commitment and self-perception. *Personality and Social Psychology Bulletin*, 22(2):133–147.

Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of verbal learning and verbal behavior*, 12(4):335–359.

DeScioli, P., Christner, J., and Kurzban, R. (2011). The omission strategy. *Psychological Science*, 22(4):442—446.

Dinner, I., Johnson, E. J., Goldstein, D. G., and Liu, K. (2011). Partitioning default effects: Why people choose not to choose. *Journal of Experimental Psychology: Applied*, 17(4):332–341.

Everett, J. A., Caviola, L., Kahane, G., Savulescu, J., and Faber, N. S. (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.

Friestad, M. and Wright, P. (1994). The persuasion knowledge model: How people cope with persuasion attempts. *The Journal of Consumer Research*, 21(1):1–31.

Gilovich, T. and Medvec, V. H. (1995). The experience of regret: What, when, and why. *Psychological Review*, 102(2):379–395.

Glaeser, E. L. (2006). Paternalism and psychology. *University of Chicago Law Review*, 73:133–156.

Gneezy, U. and Rustichini, A. (2000). A Fine is a Price. *The Journal of Legal Studies*, 29(1):1–17.

Goldstein, D. G., Johnson, E. J., Herrmann, A., and Heitmann, M. (2008). Nudge your customers toward better choices. *Harvard Business Review*, 86(12):99–105.

Grice, P. H. (1975). Logic and conversation. In Cole, P. and Morgan, J. L., editors, *Syntax and semantics, 3: Speech acts*, pages 41–58. Academic Press, New York.

Hadar, L., Sood, S., and Fox, C. R. (2011). It's not only what you know but also how knowledgeable you feel: Subjective knowledge in consumer financial decisions. Unpublished manuscript, University of California, Los Angeles.

Harrison, G. W. and List, J. A. (2004). Field experiments. *Journal of Economic literature*, 42(4):1009–1055.

Hausman, D. M. and Welch, B. (2010). Debate: To nudge or not to nudge. *Journal of Political Philosophy*, 18(1):123–136.

Heider, F. (1958). *The psychology of interpersonal relations*. John Wiley & Sons, New York.

Johnson, E. J., Bellman, S., and Lohse, G. L. (2002). Defaults, framing and privacy: Why opting in-opting out. *Marketing Letters*, 13(1):5–15.

Johnson, E. J. and Goldstein, D. G. (2003). Do defaults save lives? *Science*, 302(5649):1338–1339.

Johnson, E. J., Hershey, J., Meszaros, J., and Kunreuther, H. (1993). Framing, probability distortions, and insurance decisions. *Journal of Risk and Uncertainty*, 7(1):35–51.

Kordes-de Vaal, J. H. (1996). Intention and the omission bias: Omissions perceived as nondecisions. *Acta Psychologica*, 93(1):161–172.

Korobkin, R. (1998). Inertia and preference in contract negotiation: The psychological power of default rules and form terms. *Vand. L. Rev.*, 51:1583.

Kressel, L. M., Chapman, G. B., and Leventhal, E. (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., Tannenbaum, D., and Fox, C. R. (2017). Choice architecture 2.0: Behavioral policy as an implicit social interaction. *Behavioral Science & Policy*, 3(2):1–18.

Kullback, S. and Leibler, R. A. (1951). On information and sufficiency. *The Annals of Mathematical Statistics*, 22(1):79–86.

Larson, H. J., Cooper, L. Z., Eskola, J., Katz, S. L., and Ratzan, S. (2011). Addressing the vaccine confidence gap. *The Lancet*, 378(9790):526–535.

Larson, H. J., Jarrett, C., Eckersberger, E., Smith, D. M., and Paterson, P. (2014). Understanding vaccine hesitancy around vaccines and vaccination from a global perspective: a systematic review of published literature, 2007–2012. *Vaccine*, 32(19):2150–2159.

Lee, P. P., Stark, W. J., and Yang, J. C. (1989). Cornea donation laws in the United States. *Archives of Ophthalmology*, 107(11):1585–1589.

Lupoli, M. J., Levine, E. E., and Greenberg, A. E. (2018). Paternalistic lies. *Organizational Behavior and Human Decision Processes*, 146:31–50.

Madrian, B. C. and Shea, D. F. (2001). The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior. *Quarterly Journal of Economics*, 116(4):1149–1187.

Malle, B. F. (2004). *How the mind explains behavior: Folk explanations, meaning, and social interaction*. The MIT Press, Cambride, MA.

McAdams, R. H. and Nadler, J. (2005). Testing the focal point theory of legal compliance: The effect of third-party expression in an experimental hawk/dove game. *Journal of Empirical Legal Studies*, 2(1):87–123.

McKenzie, C. R., Sher, S., Leong, L. M., Müller-Trede, J., et al. (2018). Constructed preferences, rationality, and choice architecture.

McKenzie, C. R. M., Liersch, M. J., and Finkelstein, S. R. (2006). Recommendations implicit in policy defaults. *Psychological Science*, 17(5):414–420.

McKenzie, C. R. M. and Mikkelsen, L. (2007). A bayesian view of covariation assessment. *Cognitive Psychology*, 54(1):33–61.

McKenzie, C. R. M. and Nelson, J. D. (2003). What a speaker's choice of frame reveals: Reference points, frame selection, and framing effects. *Psychonomic Bulletin and Review*, 10(3):596–602.

Medin, D. L., Schwartz, H. C., Blok, S. V., and Birnbaum, L. A. (1999). The semantic side of decision making. *Psychonomic Bulletin and Review*, 6(4):562–569.

Mehta, S. J., Khan, T., Guerra, C., Reitz, C., McAuliffe, T., Volpp, K. G., Asch, D. A., and Doubeni, C. A. (2018). A randomized controlled trial of opt-in versus opt-out colorectal cancer screening outreach. *The American Journal of Gastroenterology*, pages 1–7.

Morris, M. W. and Larrick, R. P. (1995). When one cause casts doubt on another: A normative analysis of discounting in causal attribution. *Psychological Review*, 102(2):331–331.

Oppenheimer, D. M., Meyvis, T., and Davidenko, N. (2009). Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of Experimental Social Psychology*, 45(4):867–872.

Pichert, D. and Katsikopoulos, K. (2008). Green defaults: Information presentation and pro-environmental behaviour. *Journal of Environmental Psychology*, 28(1):63–73.

Shafir, E. and LeBoeuf, R. A. (2002). Rationality. *Annual Review of Psychology*, 53(1):491–517.

Shafir, E., Simonson, I., and Tversky, A. (1993). Reason-based choice. *Cognition*, 49(1-2):11–36.

Sher, S. and McKenzie, C. R. M. (2006). Information leakage from logically equivalent frames. *Cognition*, 101(3):467–494.

Sher, S. and McKenzie, C. R. M. (2008). Framing effects and rationality. *The Probabilistic Mind: Prospectives for Bayesian cognitive science*, page 19.

Skyrms, B. (2009). *Signals: Evolution, learning, and information*. Oxford University Press, Oxford, UK.

Spranca, M., Minsk, E., and Baron, J. (1991). Omission and commission in judgment and choice. *Journal of Experimental Social Psychology*, 27(1):76–105.

Sunstein, C. R. (2013). *Simpler: The future of government*. Simon and Schuster.

Sunstein, C. R. (2014). Choosing not to choose. *Duke Law Journal*, 64(1):1–52.

Tannenbaum, D., Fox, C. R., and Rogers, T. (2017). On the misplaced politics of behavioural policy interventions. *Nature Human Behaviour*, 1(7):0130.

Tannenbaum, D., Valasek, C. J., Knowles, E. D., and Ditto, P. H. (2013). Incentivizing wellness in the workplace: Sticks (not carrots) send stigmatizing signals. *Psychological science*, 24(8):1512–1522.

Tetlock, P. E. and Boettger, R. (1994). Accountability amplifies the status quo effect when change creates victims. *Journal of Behavioral Decision Making*, 7(4):1–23.

Thaler, R. H. and Sunstein, C. (2003). Libertarian paternalism. *American Economic Review*, 93(2):175–179.

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

Wells, G. L. and Windschitl, P. D. (1999). Stimulus sampling and social psychological experimentation. *Personality and Social Psychology Bulletin*, 25(9):1115–1125.

White, C. M., Hafenbradl, S., Hoffrage, U., Reisen, N., and Woike, J. K. (2011). Are groups more likely to defer choice than their members? *Judgment and Decision Making*, 6(3):239–251.

Young, S. D., Monin, B., and Owens, D. (2009). Opt-out testing for stigmatized diseases: A social psychological approach to understanding the potential effect of recommendations for routine HIV testing. *Health Psychology*, 28(6):675–681.