What Motivates Paternalism?

An Experimental Study

Sandro Ambuehl, B. Douglas Bernheim, and Axel Ockenfels*

August 25, 2020

Abstract

We study experimentally when, why, and how people intervene in others' choices. Choice Architects (CAs) construct opportunity sets containing bundles of time-indexed payments for Choosers. CAs frequently prevent impatient choices despite opportunities to provide advice, believing Choosers benefit. They violate common behavioral welfare criteria by removing impatient options even when all payoffs are delayed. CAs intervene not by removing options they wish they could resist when choosing for themselves (*mistakes-projective paternalism*), but rather as if they seek to align others' choices with their own aspirations (*ideals-projective paternalism*). Laboratory choices predict subjects' support for actual paternalistic policies.

*Ambuehl: University of Zurich, Department of Economics, Blüemlisalpstrasse 10, 8006 Zürich, Switzerland, sandro.ambuehl@econ.uzh.ch. Bernheim: Stanford University, Department of Economics, 579 Serra Mall, Stanford, CA, 94305, USA, bernheim@stanford.edu. Ockenfels: University of Cologne, Department of Economics, Universitätsstrasse 22a, 50923 Cologne, Germany, ockenfels@uni-koeln.de. This paper previously circulated under the title "Projective Paternalism." We thank the editor, Stefano DellaVigna, as well as three anonymous referees, Joshua Knobe, Yoram Halevy, Armin Falk, Glenn W. Harrison, Kiryl Khalmetski, Frank Schilbach, Cass Sunstein, Dmitry Taubinsky, Stefan Trautmann, and seminar and conference audiences in Berkeley, Copenhagen, Cork, Cornell, Heidelberg, Innsbruck, Nijmegen, NYU, Stanford, Toronto, and Zurich for helpful comments and discussions. Max Rainer Pascal Grossmann and Yero Samuel Ndiaye provided excellent research assistance. Ambuehl gratefully acknowledges support through a University of Toronto Connaught New Researcher Award and the University of Toronto Scarborough Wynne and Bertil Plumptre Fellowship. Parts of this paper were written while Ambuehl visited the briq Institute in Bonn. This research has been approved by the University of Toronto Research Ethics Board (protocol 10279). Ambuehl and Bernheim gratefully acknowledge funding from the Alfred P. Sloan foundation (grant number G-2017-9017). Ambuehl and Ockenfels gratefully acknowledge funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 741409); the results reflect only the authors' views; the ERC is not responsible for any use that may be made of the information it contains. Ockenfels gratefully acknowledges funding by the German Research Foundation (DFG) under Germany's Excellence Strategy (EXC 2126/1-390838866).

1 Introduction

Normative discussions of paternalism have featured in economics, philosophy, and public policy for centuries (Locke, 1764; Mill, 1869; Thaler and Sunstein, 2003). A wide range of government regulations, such as retirement savings mandates, restrictions on payday loans and investment products, various forms of consumer protection, the criminalization of suicide, and legal doctrines concerning undue inducement and unconscionability all address paternalistic concerns (Dworkin, 1971; Zamir, 1998). Paternalistically motivated social programs play a large role in the U.S. economy (Mulligan and Philipson, 2000; Moffitt, 2003; Currie and Gahvari, 2008), and in some instances even the business models of private companies are paternalistically motivated.[1] Despite the prevalence of paternalism, a *positive* investigation of the phenomenon is largely lacking. Significantly, paternalistic decision making often falls to voters (Faravelli et al., 2015), low-level government officials (Moffitt, 2006), and managers of private firms, rather than to experts, omniscient benevolent planners (Besley, 1988), or despotic autocrats (Glaeser, 2005). A large body of research catalogues the fallibility of human judgment when people choose for themselves (Kahneman, 2011). It is reasonable to assume that normative judgments that guide paternalism are also susceptible to systematic confusion or bias (Rizzo and Whitman, 2019).

This paper studies experimentally when, why, and how people act paternalistically. These questions are difficult to address in naturally occurring contexts because real-world policies with paternalistic elements generally implicate non-paternalistic concerns, such as externalities, and in any event people usually disagree about their efficacy. Conducting a laboratory experiment allows us to remove extraneous factors and study paternalistic behavior in isolation. In order to demonstrate external validity, we show that our main results extend to subjects' assessments of real-world paternalistic policies, and that these assessments are directly related to subjects' decisions within the experiment.

Subjects in the role of Choice Architects construct choice sets that determine the opportunities available to others in the role of Choosers. Each choice option consists of two payments, one received the day of the experiment ("sooner") and the other received half a year later ("later"). Impatience is costly: larger earlier payments are associated with smaller total payments. We employ the intertemporal choice domain because of its central importance across all of economics, because of *ex ante* reasons to believe that certain choices would meet with disapproval, and because it allows us to examine normative judgments that feature prominently in applications of behavioral welfare economics, including in a substantial and rapidly expanding literature (reviewed in Bernheim and Taubinsky, 2018) that argues for a wide range of paternalistic policies.

In the first part of our analysis, we show that Choice Architects frequently choose to restrict others' choice options, then we establish several fact patterns that shed light on their motivations. Significantly, subjects frequently withhold options from Choosers despite ample opportunities to provide advice. The

[1] The Vanguard Group, for instance, argues that financial advisors should attempt to help clients by "providing discipline and reason to clients who are often undisciplined and emotional" and that they "can act as emotional circuit breakers ... by circumventing their clients' tendencies" (Bennyhoff and Kinniry Jr, 2011).

typical intervention requires the Chooser to exercise a minimum level of patience. For instance, when the available options include receiving €0 sooner and €15 later, 15% of Choice Architects withhold the opportunity to receive €2 sooner and €9 later, and 31% withhold the opportunity to receive €4 sooner and €2 later. Furthermore, Choice Architects are more likely to withhold options when the relative price of current consumption is greater. Advice concerning the options that remain in Choosers' opportunity sets follows similar patterns. Thus, Choice Architects suppress impatience either by preventing it or by advising against it.

Next we ask whether these interventions are paternalistic, in the sense that Choice Architects believe they improve the Choosers' well-being (Dworkin, 1972). According to both incentive-compatible and non-incentivized measures, Choice Architects by and large believe the restrictions they impose are helpful. Furthermore, those who impose more severe restrictions believe restrictions are more beneficial. We rule out alternative motives, such as the possibility that subjects intervene simply because they prefer to do something rather than remain idle, or because they enjoy exercising control over others.

The next step in our analysis is to investigate whether Choice Architects' interventions conform to the types of normative judgments that often appear in applications of behavioral welfare economics. One common position maintains that the availability of immediate rewards, coupled with a tendency to place "too much" weight on near-term experiences ("present bias"), generate excessive impatience (e.g. O'Donoghue and Rabin, 1999, 2006; Gruber and Kőszegi, 2001, 2004; DellaVigna and Malmendier, 2004). The findings summarized above are potentially consistent with the hypothesis that Choice Architects adopt that perspective. Is the alleviation of perceived present bias in fact a central objective of the observed interventions? If it is, removing the lure of immediacy by delaying the receipt of all payments should substantially reduce interventions. Yet we find that people impose slightly *more* patience in otherwise equivalent decision problems with "front-end delay." Thus, their behavior indicates that they generally disapprove of impatience, not that they are concerned about present bias.

Another hypothesis of interest is that paternalists make decisions for others based exclusively on their own assessments of right and wrong without regard to the preferences or subjective experiences of the affected individuals. We find, on the contrary, that even non-libertarians respect Choosers' preferences to a degree, but are unwilling to tolerate sufficiently impatient choices. In particular, Choice Architects respond to information about Choosers' preferences in an accommodating direction. For instance, they impose substantially more patience on Choosers who describe themselves as impatient and unhappy than on those who claim to be content with their impatience.

The second part of our investigation asks how Choice Architects decide which options are good for others and which are bad. Their sensitivity to Choosers' inclinations and subjective experience implies that these considerations play important roles in Choice Architects' evaluations. But when their information about the Chooser is incomplete, how do they arrive at their evaluations? In *The Theory of Moral Sentiments*, Adam Smith argued that such inferences follow from self-examination: "As we have no immediate experience of

what other men feel, we can form no idea of the manner in which they are affected, but by conceiving what we ourselves should feel in the like situation." In this spirit, we identify two distinct types of departures from the benchmark case of perfect knowledge concerning others' well-being and biases, distinguished according to whether paternalists reason about others based on their own mistakes, or based on their own preferences. A *mistakes-projective paternalist* assumes others tend to share her susceptibility to error. She behaves as if she tries to help others avoid choices she herself would like to reject, but chooses nevertheless. We demonstrate that this inclination generates a *negative correlation* between the choices she makes for herself and the restrictions she imposes on others. In contrast, an *ideals-projective paternalist* behaves as if she assumes her own preferences are relevant for others, either because she thinks they tend to share her values, or because she simply believes her perspective is valid and theirs are not. Ideals-projective paternalism generates a *positive correlation* between the choices paternalists make for themselves and the restrictions they impose on others.

We find strong support for ideals-projective paternalism. More patient Choice Architects impose greater patience on others, and this pattern reflects their judgments about what the Choosers ought to do, rather than a greater inclination to intervene. They also believe more strongly that restrictions enhance the Choosers' well-being.

Next we ask whether paternalists simply express their welfare judgments based on accurate beliefs about others, or whether paternalistic interventions are susceptible to systematic cognitive biases. To answer this question, we focus on Choice Architects' beliefs about the choices unrestricted Choosers would make. We examine how those beliefs relate not only to the restrictions Choice Architects impose, but also to their own preferences. We find that interventions are closely related to beliefs about Choosers' inclinations. Moreover, those beliefs are systematically biased: subjects assume (incorrectly) that the choices of others tend to resemble their own (a *false consensus effect*, Ross et al., 1977). As a result, despite the tendency for more patient Choice Architects to impose greater patience, patient and impatient Choice Architects expect their mandates to bind with about the same frequency. We conclude that interventions are systematically misguided, even according to the Choice Architects' aims.

The third part of our analysis explores the external validity of our findings. We elicit subjects' support for imposing sin taxes and for regulation of high-interest, short-term lending in a neighboring country. We also measure subjects' own consumption of the targeted products. We find that real-world preferences for paternalistic policies that only impact others are strongly correlated with Choice Architects' decisions in the behavioral portion of our experiment. Additionally, the relationships between policy preferences and own consumption are consistent with ideals-projective paternalism. For instance, lighter drinkers express significantly more support for an increase in alcohol taxes in another jurisdiction.

Our work is related to a small empirical literature on paternalistic behavior. The closest parallels are Uhl (2011) and Krawczyk and Wozny (2017). In both studies, subjects choose between one of two options for

themselves and decide whether to eliminate one of the options for others.[2] While these studies examine the proclivity to intervene, as well as correlations between chosen interventions and the options subjects select for themselves, they do so in settings that preclude the dispensation of advice. Accordingly, it is impossible to tell whether the subjects see intervention as intrinsically desirable, or merely as the only feasible method of expressing their opinions. Furthermore, the close juxtaposition of essentially identical decisions (for the subjects themselves and for others) may well introduce spurious correlation through anchoring and/or a demand for consistency, which undermines inferences concerning the connection between own-preferences and interventions. Setting these important design issues aside, these papers have more limited objectives than ours, and consequently do not address the other important questions that structure our analysis: whether interventions are benevolently motivated; whether paternalists care about and respond to information concerning the preferences of the impacted parties; whether they exhibit less inclination to intervene when those parties make their choices under conditions that are commonly thought to mitigate decision biases; whether paternalistic interventions implicate systematic cognitive biases; and whether paternalism in the lab relates to real-world policy preferences. Likewise, they do not address the distinction, central to this paper, between ideals-projective and mistakes-projective paternalism. In a more recent paper, Bartling et al. (2020) focus on decisions to intervene paternalistically by removing dominated options. They largely abstract from motivations involving disapproval of others' preferences, which is a central feature of the current paper.

Several other lines of work bear on the empirical analysis of paternalism. One examines the tendency for professional advisors to steer their clients toward the same options they choose for themselves (Foerster et al., 2017; Linnainmaa et al., forthcoming) – a possible reflection of ideals-projective paternalism. A second studies how people feel about being in situations where others can influence or constrain their choices (Fehr et al., 2013; Bartling et al., 2014; Kataria et al., 2014; Lübbecke and Schnedler, 2018; Ackfeld and Ockenfels, 2020).[3] A third studies social disapproval of ostensibly *repugnant transactions* (Roth, 2007), such as paid organ donation (Basu, 2003, 2007; Leider and Roth, 2010; Elias et al., 2015a,b, 2019; Ambuehl, 2017; Ambuehl and Ockenfels, 2017; Clemens, 2018; Exley and Kessler, 2017).[4] A fourth explores how people make surrogate choices for others in settings where the surrogate cannot leave the affected individual with flexibility (see Ifcher and Zarghamee, 2020, for a review).

The empirical study of paternalism connects with various other branches of the literature. The literature on *libertarian paternalism* argues that authorities can use *nudges*, rather than coercion, to adjust behavior in directions the authorities (or their policy analysts) deem beneficial (Thaler and Sunstein, 2003; see Loewenstein and Haisley, 2007; Benartzi et al., 2017, for reviews). Our analysis sheds light on the types of paternalistic judgments that motivate nudges, and it identifies common biases that nudge designers could in

[2] In Krawczyk and Wozny (2017), the options are two lunch items, one healthy, the other unhealthy. In Uhl (2011), the options are whether to make a second choice (i.e., choosing the point in time at which to collect a rising payment) either in advance or in the moment. The empirical literature on paternalistic behavior also includes Jacobsson et al. (2007); Gangadharan et al. (2015); Lusk et al. (2013); Schroeder et al. (2017).

[3] A larger theoretical literature relating to paternalism includes Saint-Paul (2004); Carlin et al. (2013); Bisin et al. (2015); Altmann et al. (2017); Laibson (2018).

[4] Ambuehl et al. (2015) sketch a model with agents who exhibit a form of ideals-projective paternalism that explains why the introduction of monetary incentives can cause people to judge transactions as unethical (e.g. paid kidney donation), even though they approve of the same transaction in the absence of such incentives (e.g. in-kind kidney exchange).

principle learn to avoid. We contribute more broadly to a literature in cognitive science on moral heuristics (see Sunstein, 2005; Gigerenzer, 2008, for reviews). There is also a connection to a literature on *projection bias* (Van Boven et al., 2000; Loewenstein et al., 2003; Madarasz, 2012), in that we exhibit a manifestation of this phenomenon in the normative domain. Finally, a collection of recent empirical studies evaluate particular paternalistic policies by deploying normative principles from behavioral welfare economics; see Bernheim and Taubinsky (2018) for a review, and Bernheim and Rangel (2009) or Bernheim (2016) for foundations.

The remainder of this paper proceeds as follows. Section 2 outlines our experimental design. Section 3 demonstrates that Choice Architects frequently force others to choose patiently, and that they believe these interventions benefit the Choosers. It also explores the responsiveness of paternalistic interventions to the introduction of front-end delay, and of information about Choosers' preferences. In Section 4, we investigate hypotheses about the formation of Choice Architects' views concerning Choosers. We begin by formalizing the paternalists' decision problem, introducing the concepts of ideals-projective and mistakes-projective paternalism, and deriving their implications. After documenting patterns that point to ideals-projective paternalism, we explore the prevalence and implications of mistaken beliefs about the selections Choosers would make in absence of interventions. Section 5 demonstrates that laboratory choices predict support for real-world paternalistic policies, and it documents ideals-projective paternalism regarding those policies. Finally, Section 6 outlines directions for further research.

2 Experiment Design

We begin by explaining the design of our experiment. Section 2.1 describes the main types of decision problems we use to investigate paternalism. Section 2.2 provides an overview of the structure of the experiment. The remaining subsections then present details concerning incentivization (Section 2.3), the Choosers' decisions (Section 2.4), and implementation (Section 2.5). For easier readability, this section condenses the presentation of our design. A comprehensive description appears in Appendix D.1.

2.1 Main elements of the experiment

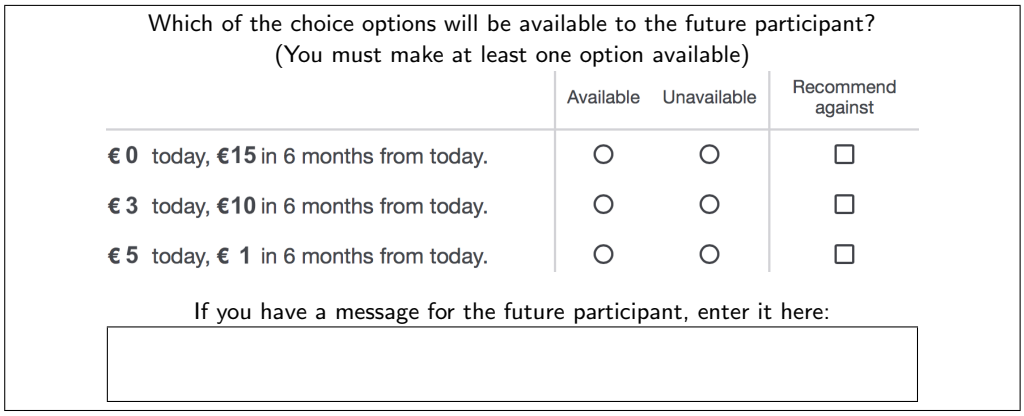
Each subject in our experiment is either a Choice Architect or a Chooser. Our interest is in the Choice Architects, who determine the set of options that will be available to Choosers. The experiment's main elements, described below, shed light on the following four questions: (i) How do Choice Architects construct opportunity sets for Choosers? (ii) Do Choice Architects believe that withholding options helps or hurts Choosers? (iii) What options do Choice Architects select for themselves? (iv) What do they believe Choosers would select absent restrictions? We discuss each of these elements in turn.

Construction of opportunity sets The Choice Architect constructs the Chooser's opportunity set from a menu of three options, as illustrated in Panel A of Figure 1. Each option is a bundle of two monetary payments, one received the day of the experiment, the other received with a half-year delay. We design the options so that a Chooser can increase his present payment only by accepting a smaller amount overall.

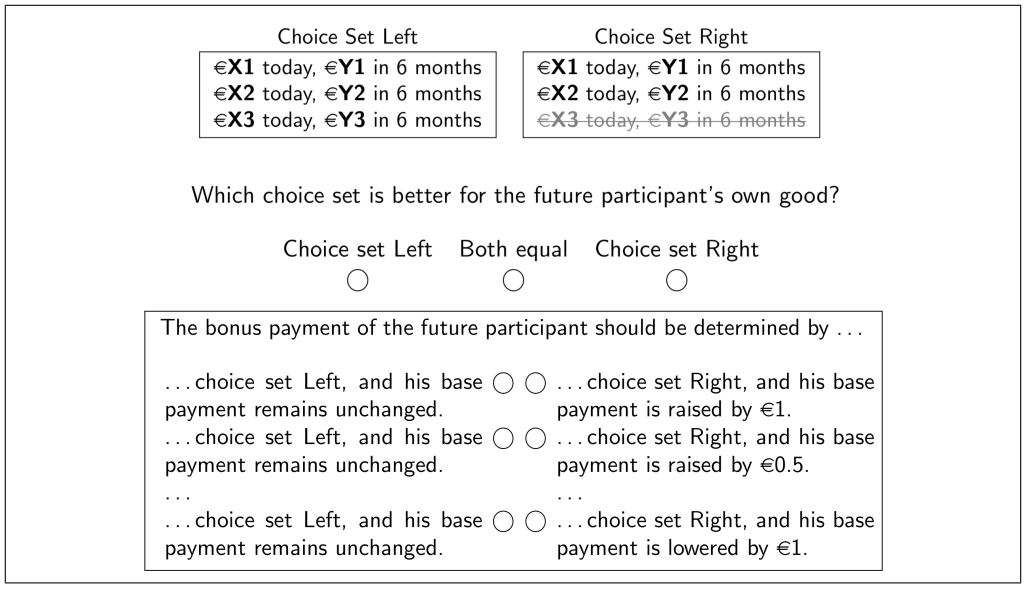
The Choice Architect must *actively* decide whether the opportunity set will include each option; neither inclusion nor exclusion is a default. The sole restriction is that each opportunity set must include at least one option. We emphasize to subjects that there are no right or wrong answers, and that they should construct opportunity sets based on their genuine views.

Figure 1: Decision screens for the Choice Architect in the Main condition.

A. Constructing the Chooser's opportunity set.



B. Beliefs about effect on Chooser's well-being.



We study the characteristics of the opportunity sets Choice Architects construct. Three design features limit the possible explanations for imposing restrictions. First, Choice Architects can advise Choosers. Specifically, in each round, the Choice Architect can write an unrestricted message to the Chooser, which the Chooser observes before making his decision. The Choice Architect can also convey disapproval of any option by clicking a button. In that case, the Chooser sees a red asterisk next to the corresponding option, accompanied by a statement that a previous participant advises against it. The existence of these opportunities is central to our objective. Without them, Choice Architects might remove options simply

because they have no other way to convey opinions and advice, rather than because they perceive a genuine need for restrictions. Accordingly, this feature allows us to detect situations in which Choice Architects feel they cannot trust informed Choosers to make good decisions (an expression of paternalism), rather than situations in which Choice Architects merely believe they have better information than Choosers (a state of affairs that does not necessarily implicate paternalism).

Second, we ensure that Choice Architects can only influence Choosers' outcomes, not their decision processes. In particular, the Choice Architect cannot save the Chooser time or effort, or spare him the ordeal of resisting temptation. The reason is that the Chooser ranks all three of the options that might be in his opportunity set, without knowing which are actually available or how their availability is determined. He then receives the option he has ranked most highly among those that are actually available. Choice Architects are aware of this procedure.

Third, the Choice Architect's material payoff is independent of her decisions concerning the Chooser's opportunity sets. Accordingly, the Choice Architect's only plausible motivations involve her feelings about the Chooser's autonomy or consequences. Here, our objective is to avoid confounding paternalistic motivations with conflicts of interest. This feature mirrors an important property of many paternalistic decisions. For example, members of Internal Review Boards charged with protecting human subjects are usually precluded from having personal interests in any research that is subject to their oversight.

Three additional features of this setting merit emphasis. First, our experiment focuses on *hard paternalism* (restricting opportunity sets) rather than *soft paternalism* (influencing choice without changing opportunity sets). Soft paternalism introduces other potentially confounding considerations. For example, the attractiveness of employing a nudge depends in part on beliefs about its efficacy. In our setting, efficacy is unambiguous. Second, we study paternalistic decisions by individuals, rather than groups. While many paternalistic policies result from group decision making (e.g., through voting), the judgments of individuals are always central. Third, by using opportunity sets involving bundles of immediate and delayed monetary payments, our experiment introduces a plausible *a priori* rationale for paternalism: people commonly view patience as virtuous and impatience as reflecting weakness.

Elicitation of beliefs about the welfare effects of restrictions A Choice Architect's decision to withhold options is paternalistic only if she believes it promotes the Chooser's well-being (Dworkin, 1972). We measure these beliefs in two ways on a single screen, as illustrated in Panel B of Figure 1. The Choice Architect sees an opportunity set on the left, along with a subset of the same options on the right. As explained subsequently, in some rounds the subset on the right is the one the Choice Architect constructed; in other rounds it is given exogenously.[5]

Choice Architects first answer a simple non-incentivized question: *Which opportunity set is better for the future participant?* They select between *Opportunity Set Left*, *Both equal*, and *Opportunity Set Right*. Second,

[5] In most cases, the opportunity set on the left contains three options. However, for half of the rounds involving exogenous restrictions, the opportunity set on the left contains two options, and the one on the right contains only the most patient option. See Section 3.5 and Appendix B.1 for additional details.

Choice Architects complete a decision list that elicits their beliefs about the welfare effects of restrictions. Assuming they are not entirely indifferent towards Choosers, the elicitation is incentive-compatible. Each line of the list presents a binary choice of the following form: *The payment of the future participant should be determined by ... (i) the opportunity set on the left, and the participant's completion payment will remain unchanged, OR (ii) the opportunity set on the right, and the participant's completion payment will be raised / lowered by* €p, with $p \in \{1, 0.5, 0.3, 0.1, 0.05, 0, -0.05, -0.1, -0.3, -0.5, -1\}$. A Choice Architect may believe, for instance, that enlarging an opportunity set requires positive compensation of €0.4 to the Chooser if the additional choice options create opportunities for errors. A benevolent Choice Architect will prefer the first option if $p = -0.5$, but will prefer the second if $p = -0.3$. Thus, the transfer p at which a benevolent Choice Architect switches from *(i)* to *(ii)* reveals her beliefs about the payment that compensates the Chooser for receiving the opportunity set on the right over the one on the left. We impose no restrictions on how subjects fill in these lists such as monotonicity. Because we implement (at most) one decision from one list, the Choice Architect has an incentive to choose in accordance with her genuine preferences, regardless whether she is benevolent or malevolent.[6]

Elicitation of Choice Architects' time preferences Our investigation of paternalistic mechanisms involves comparisons between the opportunity sets Choice Architects construct for Choosers and the choices they make for themselves. All Choice Architects complete six decision lists such as the one shown in Figure 2. Each line is a choice between €x_{early} the day of the experiment and €x_{late} t months later.[7]

Figure 2: Choice Architects' own intertemporal choices.

On each line, choose the option you genuinely prefer:

€8 on day of experiment	☐	☐	€10 in 1 month after the experiment
€8 on day of experiment	☐	☐	€10 in 2 months after the experiment
€8 on day of experiment	☐	☐	€10 in 3 months after the experiment
€8 on day of experiment	☐	☐	€10 in 4 months after the experiment
€8 on day of experiment	☐	☐	€10 in 5 months after the experiment
€8 on day of experiment	☐	☐	€10 in 6 months after the experiment

To ensure that anchoring or a demand for consistency does not produce an artificial relationship between the decisions Choice Architects make for themselves and those that impact others, we make it difficult for them to compare options across these settings. When choosing for themselves, the amount of delay differs across the options but the monetary payments are the same. When constructing opportunity sets for Choosers, the monetary payments differ across options but the amount of delay is the same. We also limit the potential influence of confounding mechanisms by requiring Choice Architects to make choices for themselves online three to six days before the laboratory session, and by interspersing these tasks with decision lists involving risk taking to obfuscate their purpose.

[6] Our elicitation differs in two ways from the types of incentive-compatible belief elicitation techniques commonly used in experimental economics. First, we add or subtract the amount p to or from the Chooser's completion payment, not the Choice Architect's completion payment, even though the Choice Architect is the party expressing the beliefs. Second, we do not compensate the Choice Architects based on the distance between their expressed beliefs and some objective truth. Our approach more closely resembles the incentive-compatible methods used to elicit willingness-to-pay rather than those used to elicit beliefs.

[7] We use $(x_{early}, x_{late}) \in \{(2, 10), (5, 10), (8, 10), (2, 15), (7, 15), (12, 15)\}$ and t between 1 and 6 months in steps of one month.

The temporal separation between the online session and the lab session implies that the earlier payments in the online session are not immediate. This lack of immediacy could in principle cause Choice Architects to think about the options available for themselves in the online session and those available to Choosers in the laboratory session somewhat differently. Consequently, towards the end of the laboratory session, Choice Architects make additional selections for themselves from each three-option set for which they previously constructed Choosers' opportunity sets. While these choices potentially raise greater concerns about anchoring or a demand for consistency than those made in the online session, we find little evidence that these confounds are in fact present in our experiment; see Appendix Section B.2.

Elicitation of beliefs about Choosers' unrestricted choices We elicit Choice Architects' beliefs about the distributions of unrestricted choices made by ten previous subjects for each menu encountered elsewhere in the experiment. Choice Architects drag and drop ten tags labeled "Participant" into bins representing each of the three choice options. Systematic discrepancies between these beliefs and the actual (known) distribution of unrestricted choices may imply that paternalistic interventions are predicated on mistaken perceptions of Choosers' tendencies. To ensure incentive compatibility, a Choice Architects' compensation may be based entirely on the accuracy of their reported beliefs.

2.2 Structure of the experiment and additional elicitations

As noted above, the experiment consists of an online component and a laboratory component. The online component elicits Choice Architects' own intertemporal preferences. The laboratory component consists of three stages. Stage 1 includes 14 rounds of paternalistic decisions and assessments of perceived welfare effects that employ various menus of options from which Choice Architects construct choice sets. The *Main* condition comprises four rounds, each of which proceeds as described above. These rounds employ menus 1 to 4 shown in Table 1. For each subject, we randomly select the round involving either menu 1 or 2 and delay both the early and the late payment for each option in that menu by one week. As we discuss subsequently, the introduction of front-end delay allows us to test some specific ideas about the nature of paternalism. Each Choice Architect also participates in three additional conditions (ten rounds in total) for which we alter the decision problems described above to test specific hypotheses about the determinants of paternalism. These include the *Exogenous Restriction* condition (four rounds using menu 6), the *Chooser Information* condition (four rounds using menu 5), and the *Induced Chooser Preference* condition (two rounds using menus 3 and 4). We describe these conditions when we encounter the hypotheses they address in Sections 3.2, 3.4, and 3.5, respectively.

Each subject proceeds through the rounds of Stage 1 in an individually randomized order. In each round, the Choice Architect first constructs the Chooser's opportunity set, and then reveals her beliefs about whether the three-option opportunity set or a subset thereof is better for the Chooser. In the Main condition, the comparator subset is the one the Choice Architect constructed herself.

In Stage 2, we collect additional data that shed light on aspects of the Stage 1 decisions. First, Choice

Table 1: Menus of options from which Choice Architects construct opportunity sets.

	Menu 1		Menu 2		Menu 3	
Option	Today	In 6 months	Today	In 6 months	Today	In 6 months
Most patient	\$0	\$15	\$0	\$15	\$0	\$15
Middle	\$3	\$10	\$3	\$9	\$2	\$12
Least patient	\$5	\$1	\$6	\$1	\$4	\$2

	Menu 4		Menu 5		Menu 6	
Option	Today	in 6 months	Today	In 6 months	Today	In 6 months
Most patient	\$0	\$15	\$0	\$15	\$0	\$15
Middle	\$2	\$9	\$4	\$6	\$3	\$7
Least patient	\$4	\$2	\$5	\$1	\$4	\$1

Notes: Menus 1 to 4 correspond to the Main condition. Menu 5 corresponds to the Chooser Information condition (section 3.4). Menu 6 corresponds to the Exogenous Removal condition (section 3.5).

Architects make surrogate choices for Choosers. These tasks are identical to those in Stage 1, except we require Choice Architects to select a single option. Assuming Choice Architects are benevolent, these decisions reveal the options they deem best for Choosers. Unlike decisions to restrict opportunity sets, surrogate choices do not implicate Choice Architects' willingness to intervene. Second, to incentivize attentiveness, we administer an eight-question test on specific features of the experiment. We tell subjects about this test in advance, and advise them that their performance on it could completely determine their earnings.[8] Third, we elicit beliefs about previous Choosers' unrestricted choices, as described above.[9] Fourth, Choice Architects select an option for themselves from each menu, also as described above. Fifth, we ask Choice Architects to adjust the completion payment of a different Chooser. They can costlessly increase that payment by €1, leave it unchanged, or decrease it by €1. We use their responses to gauge whether they are benevolent or spiteful toward Choosers.

In Stage 3, subjects express opinions about four real-world paternalistic policy proposals, and then provide information about their own inclinations to engage in the affected activities. We use this information to evaluate the generalizability of our findings. The experiment ends with a brief memory check concerning choices subjects made in the Online component.

Appendix D.1 presents comprehensive detail about all design elements.

2.3 Incentives

Choice Architects' decisions concerning Choosers. There is a 25% chance that we match any given Choice Architect with a Chooser. For each match, we randomly draw one of the rounds in which the Choice Architect makes a decision concerning the Chooser. With 50% probability, we implement the decision from the first half of that round. With the remaining 50% probability, we determine the Chooser's opportunity set

[8] Appendix B.3 lists the test questions and frequencies of correct responses. Subjects do not learn anything about the content or focus of the test before making decisions concerning the Chooser. The test consists of eight questions about the stimuli the Choice Architects encountered. It does not refer to Choice Architects' own decisions.

[9] Half of our subjects, chosen at random, complete these elicitations before Stage 1. We control for this order in all of our regressions.

by randomly drawing a line from the decision list in the second half of the round (which elicits the perceived benefit of a restriction) and implementing the Choice Architect's selection. Separately, with 25% probability, we implement the Choice Architect's chosen adjustment to the completion payment for a different Chooser who we assign at random. Choice Architects know that Choosers will participate in a subsequent laboratory session. We inform Choice Architects of the matching and implementation probabilities, and explain that, if they are matched to a Chooser, no other subject will influence the Chooser's opportunities. We also let Choice Architects know that no other subject has manipulated the choice problems that determine their own payments.

Choice Architects' own payment. A Choice Architect's payment is determined either by the online component or by one of the following: the attention test, elicitation of beliefs concerning Choosers' unrestricted selections, or choices made for themselves. Each of these four alternatives is equally likely.[10] For the elicitation of beliefs concerning unrestricted Choosers' selections, the Choice Architect receives €10 minus the number of tags we must reassign to make the elicited and observed distributions of choices coincide.[11] For the attention test, the Choice Architect receives €1 for each correct answer. For the online component or the Choice Architects' decisions for herself in the laboratory component, we implement one randomly selected decision within the corresponding task block. Each Choice Architect also receives a show-up payment of €4.5 and a completion payment of €8.[12]

2.4 Choosers

Choosers participate after all Choice Architect sessions are completed. Each Chooser first selects one of four self-descriptive statements (see Section 3.4) and then ranks the elements of one menu. He receives the option he ranks highest among those his matched Choice Architect makes available. Each Chooser also receives a show-up payment of €4.5 and a completion payment of €8. If the Choice Architect task selected for implementation involves the incentive-compatible elicitation of beliefs concerning welfare effects, we raise or lower the corresponding Chooser's completion payment based on the Choice Architect's decision.

2.5 Implementation

We conducted 16 sessions with a total of 303 Choice Architects during the summer of 2018 at the Cologne Laboratory for Economic Research. Each lasted approximately 90 minutes. We recruited 100 additional subjects to study other hypotheses (see the discussion of the Choice Distribution Information condition in Appendix B.4). Separately, 124 subjects participated as Choosers (see Appendix B.5).

The experiment is computer-based. We display instructions on-screen, and intersperse comprehension

[10] Subjects learn at the beginning of the online component that there is a 25% chance a single decision from that component will determine their payment, and a 75% chance the laboratory component will determine it, but they do not learn at the outset what the latter possibility will entail.

[11] Our elicitation procedure is the balls-in-bins method described in Delavande et al. (2011). Truthful revelation is optimal for a risk-neutral subject. Subjects understand this scheme more easily than alternatives. While risk aversion theoretically generates a tendency for measured beliefs to be overly dispersed, Choice Architects' risk preferences, elicited in the online component, predict neither the location nor the dispersion of elicited belief distributions.

[12] The completion payment in the first two sessions was €5, which we increased after feedback that subjects perceived the study payment to be low.

checks which subjects must complete correctly to continue. Those failing the comprehension checks can review the instructions and try again (all subjects eventually passed). The comprehension checks emphasize that there are no right or wrong answers for decisions affecting the Choosers. We process all incentive payments through PayPal and cover all transaction fees.

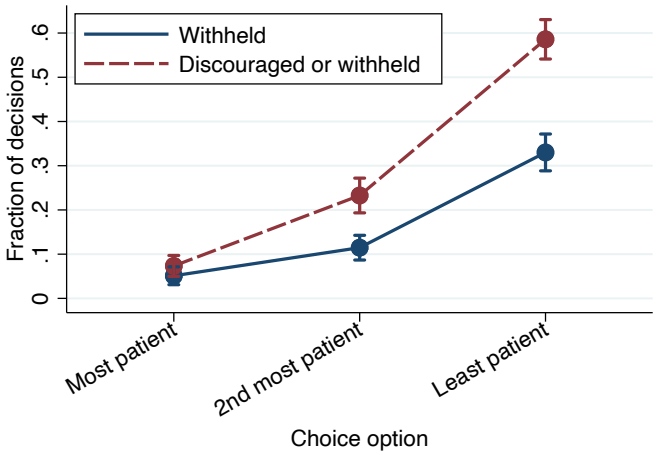
3 Properties of paternalistic interventions

3.1 Do people intervene and, if so, in what way?

The first step in our analysis is to ask whether people intervene into others' decisions and to characterize the nature of those interventions. We find that interventions are common, and that they generally limit impatient choices, especially when impatience is costly.

Figure 3 shows the frequencies with which Choice Architects make specific types of options unavailable, averaged across rounds in the Main condition (excluding those with front-end delay). The prevailing tendency is for Choice Architects to prevent impatient choice. They remove the least patient option 32.7% of the time, the middle option 11.6% of the time, and the most patient option 5.1% of the time. When we limit attention to the 86.3% of Choice Architects who are altruistic (in the sense that they costlessly increase a second Chooser's completion payment in Stage 2), the frequency with which they remove the most patient option drops by half (to 2.5%). The removal frequency for the middle option falls slightly (to 10.6%), and there is no change for the least patient option. As Table 2 shows, the pattern of withholding impatient options rather than patient ones emerges for all four menus.

Figure 3: Withheld and discouraged options.



Notes: The solid line displays the fraction of cases for which the Choice Architect withholds options, categorized by relative patience, from the Chooser. The vertical distance between the solid and dashed lines represents the fraction of cases in which the Choice Architect advises against the option but does not withhold it. We average the data across the menus (1 to 4) in the Main condition, excluding rounds with front-end delay. Whiskers indicate 95%-confidence intervals based on estimates of standard errors clustered at the subject level.

Imposing a minimum degree of patience is also the modal behavioral pattern on the individual level: 44.9% of our subjects remove at least one option from at least one of the opportunity sets, and never remove an option without also excluding less patient options. Choice Architects who never remove any option in the

Main condition comprise the second largest category (38.3%). We refer to this group as Libertarians. A small fraction (4.3%) of Choice Architects impose an upper bound on patience.[13] The remaining 12.5% of Choice Architects impose non-monotonic restrictions.

Choice Architects' decisions depend not only on the rank (by patience) of an option within a menu, but also on the cost of impatience, defined as the total amount of money the Chooser forfeits by acting impatiently. In menu 3, for instance, the Chooser sacrifices a total of €1 (€15 − €14) by selecting the middle option rather than the most patient option. In menus 1, 2, and 4, the corresponding figures are €2, €3, and €4, respectively. Removal rates for the middle option increase monotonically with the forfeited amount: they are 8.3% for menu 3, 9.3% for menu 1, 14.4% for menu 2, and 14.5% for menu 4 (see Table 2). A similar pattern holds for the least patient option. By selecting the least patient option from menu 2 rather than the most patient option, a Chooser forfeits €8 in total, whereas the corresponding amount is €9 for each of the remaining menus. The removal rate for the least patient option is 23.5% for menu 2, compared with 32%, 38.9%, and 31.4% for menus 1, 3, and 4, respectively. To quantify these relationships, we estimate an OLS regression relating the availability of an option to the total amount of money a Chooser would forfeit by selecting it rather than the most patient option, controlling for the rank of the option within the choice set, as well as session and order fixed effects. We find that, for each additional Euro sacrificed when selecting an impatient option, the frequency with which Choice Architects withhold that option increases by 2.9 percentage points ($p < 0.01$, see Appendix B.6 for details).

Table 2: Fraction of options withheld in Main condition.

Menus	Percent option unavailable			Observations	Subjects
	Most patient	Middle	Least patient		
1, 2, 3, 4	5.1	11.6	32.7	909	303
1	6.0	9.3	32.0	150	150
2	5.2	14.4	23.5	153	153
3	4.3	8.3	38.9	303	303
4	5.3	14.5	31.4	303	303

Notes: Decisions with front-end delay are excluded.

The frequency with which subjects advise against options but do not withhold them exhibits similar patterns; see Figure 3. Subjects advise against impatient options substantially more often than against patient ones, just as they tend to remove the former more than the latter. Thus, patterns of hard paternalism (prohibitions) and soft paternalism (advice) mirror each other. Significantly, the pattern of advice is essentially the same, both qualitatively and quantitatively, for subjects who never withhold options. Subjects write free-form messages much less frequently (only 21.5% of subjects ever send a message exceeding 5 characters in length), but when they do, they typically articulate paternalistic perspectives concerning the desirability of patience. Examples include "*Patience is a virtue*," "*Stay rational. Don't weigh the present more than the future*," and "*It is important to practice patience.*"

[13] These subjects remove at least one option from at least one choice set, and never remove an option without also excluding more patient options.

3.2 Do people believe their interventions are beneficial?

Having shown that interventions occur in our experiment, we now ask whether they are paternalistic. A restriction on free choice is deemed paternalistic only if the party responsible for it believes that it benefits those it affects (Dworkin, 1972). In this section, we show that our Choice Architects think the restrictions they impose are beneficial, and that those who impose stricter mandates believe the benefits are greater. Accordingly, we are indeed observing paternalism.

Table 3: Subjects' beliefs about the welfare effects of withholding options.

Dependent variable	(1) Belief smaller opportunity set better for Chooser	(2)	(3) Negative of Compensating Variation	(4)
Smaller set	Chosen	Exogenous	Chosen	Exogenous
Menus	1 - 4 without FED	6	1 - 4 without FED	6
	Summary statistics for dependent variables			
Distribution of beliefs				
Evaluation opportunity set is				
better	0.296	0.345		
same	0.600	0.107		
worse	0.105	0.548		
Mean negative CV			0.021	-0.224
Distribution of beliefs if options withheld				
Evaluation opportunity set is				
better	0.652	0.455		
same	0.091	0.118		
worse	0.257	0.428		
Mean negative CV			0.051	-0.122
	Regression results			
1 option withheld	0.320***		0.019	
	(0.068)		(0.050)	
2 options withheld	0.496***		0.113	
	(0.104)		(0.079)	
Mean # options withheld		0.686***		0.251***
in Main condition		(0.073)		(0.055)
Comparison set contains		-0.182**		-0.253***
most patient option only		(0.088)		(0.062)
Observations	909	606	866	574
Number of subjects	303	303	297	292

Notes: The top half of the table provides summary statistics for the dependent variables, and the bottom half presents regression results. FED stands for front-end delay. Unit of observation: subject-round pair. Method: Columns 1 and 2, OLS; columns 3 and 4, interval regression. Dependent variables: The dependent variable for columns 1 and 2 measures whether the Choice Architect considers the smaller opportunity set better, equally good, or worse for the Chooser than the unrestricted set, coded as 1, 0, and -1, respectively. For columns 3 and 4, it is the negative of the Choice Architect's beliefs about the compensating variation (CV) of reducing the opportunity set. For all columns 1 and 3, the smaller opportunity set is the one the Choice Architect has constructed; for columns 2 and 4, we construct it by exogenously removing either the least patient option, or the least patient and middle options. Controls: All regressions control for altruism and spite, and include session, and order fixed effects. Columns 1 and 3 also control for menu fixed effects Samples: Columns employ data from different conditions, as indicated. For columns 1 and 3, the row *Distribution of beliefs if some option withheld* includes all observations in which the Choice Architect withheld one or more of the options in rounds of the Main condition without front-end delay. For columns 2 and 4, that row includes observations from all rounds in the Exogenous Restriction condition for Choice Architects who withheld at least one option in the Main condition. In the bottom panel, each column corresponds to a separate regression. Columns 3 and 4 exclude subjects with multiple switches in the elicitation of compensating variations. Standard errors: clustered by subject. $^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$.

We begin by analyzing subjects' responses to the question of whether the full opportunity set or the opportunity set they have constructed is better for the Chooser. Column 1 of Table 3 (top half) shows the distribution of responses in the Main condition (without front-end delay). In 30% of cases, subjects report that the choice set they have constructed is better for the Chooser than the unrestricted set, while 11% report that it is worse. These frequencies are artificially attenuated by the presence of libertarian subjects who never remove options, and who therefore largely report that neither set is better. For rounds in which subjects imposed restrictions, 65% describe the restricted set as better, while 25% describe it as worse. Overall, subjects are far more likely to believe their interventions are helpful than hurtful.

To evaluate the statistical significance of these findings, we regress the subject's report about the welfare effect on indicators for the number of options removed, encoding the responses "better," "same," and "worse" as +1, 0, and -1, respectively (Appendix B.7 reports the corresponding ordered probit regressions). We use the full sample and control for altruism versus spite on the part of the Choice Architect, i.e., whether she has chosen to costlessly increase or decrease a second Choosers' completion payment by €1. We include session, order, and menu fixed effects, and cluster standard errors by subject. On average, a Choice Architect is more likely to believe that her action benefits the Chooser when she removes one option rather than none, and the difference is highly statistically significant, as the estimates in column 1 show. When the Choice Architect removes two options, the estimated coefficient is even larger, but the increment is only marginally significant ($p = 0.11$).[14]

One purpose of the *Exogenous Restrictions* condition, which we mentioned briefly in Section 2.2, is to allow us to study beliefs about the welfare effects of prespecified restrictions whether or not Choice Architects select them. Here we examine two of the four rounds in this condition (see Section 3.5 for an explanation of the other two rounds). There are two main differences between the welfare evaluations in these rounds and those in the Main condition. First, they involve comparisons between the unrestricted set and an exogenously specified subset that includes either the most patient option or the two most patient options. The number of items in the comparison set (one or two) is determined randomly for each subject and is the same in the both rounds. Second, this condition employs menu 6 (see Table 1).[15]

Because exogenously imposed restrictions need not coincide with the interventions Choice Architects deem best for Choosers, we would expect to find less optimism about their welfare effects. Indeed, as shown in top half of Table 3, column 2, Choice Architects say that the restrictions make Choosers better off 35% of the time and worse off 55% of the time. Of course, these figures include libertarians. Non-libertarian Choice Architects (who, by definition, withhold some option in the Main treatment) say that the restrictions make Choosers better off 47% of the time and worse off 42% of the time. Moreover, even non-libertarians may view

[14] We also find that spiteful Choice Architects believe much less strongly that their restrictions benefit Choosers. We caution, however, that this estimate is based on only nine spiteful subjects.

[15] For the purpose of examining Choice Architects' beliefs about welfare effects, the two rounds are essentially identical, so it is especially important to report standard errors clustered at the subject level reported in Table 3. The distinction between these rounds is that we provide the Choice Architect with a different default when asking her to construct the Chooser's opportunity set: in one round, the default consists of all three options, and the Choice Architect can remove either the least patient option or the least patient and middle options; in the other round, the default consists of the most patient option, and the Choice Architect can add either the middle option or the middle and least patient options. We examine Choice Architects' selected opportunity sets for these rounds, as well as for the other two rounds of the Exogenous Restrictions condition, in Section 3.5.

the removal of two options as excessive; indeed, in the Main treatment, they remove 0.86 options per round on average. Non-libertarian Choice Architects say that removing the most impatient option, and nothing else, makes the Chooser better off 52% of the time and worse off 35% of the time (not shown in the table). On average, these subjects clearly believe that a minimal restriction helps: the difference in frequencies is 19 percentage points ($p = 0.078$).

We report associated regression results in the bottom half of Table 3, column 2. We use OLS as a convenient method for calculating conditional means, but do not interpret the estimates as causal effects. The regression relates welfare evaluations of exogenous restrictions, calibrated as before, to a proxy for the Choice Architect's proclivity to intervene (the average number of options removed in the Main condition), and a dummy variable indicating that the comparison set is limited to the most patient item. Once again, we estimate the regression with the full sample, control for altruism versus spite on the part of the Choice Architect, and include session and order effects. Two conclusions follow from this regression. First, Choice Architects who have demonstrated a greater proclivity to restrict Choosers in the Main condition are substantially more likely to believe that exogenously specified restrictions are beneficial. Second, Choice Architects are more likely to believe that a restriction is beneficial if it removes one item rather than two.

Columns 5 and 6 replicate columns 1 and 2 using dependent variables based on our second measure of welfare effects: beliefs about compensating variations, elicited through an incentive-compatible procedure. Here we estimate interval regressions, including only those subjects whose choices in the multiple-decision lists are consistent with a preference for increasing the Chooser's payment.[16] Results for evaluations of exogenous restrictions corroborate our findings based on unincentivized evaluations (compare column 6 to column 2). Naturally, magnitudes differ because the scales of the dependent variables are not comparable. Results for evaluations of chosen restrictions are consistent with our findings based on unincentivized evaluations (compare column 5 to column 1), but the coefficients are not statistically distinguishable from zero at conventional levels of confidence. This loss of statistical precision may reflect the complexity of the incentive compatible elicitation, which could introduce noise. While this evidence is correlational, Section 3.4 shows that exogenously varying information about Choosers alters beliefs about welfare effects. Moreover, Section 3.5 provides evidence against the hypothesis that causality flows from choices to beliefs through a rationalization mechanism.

Subjects' own explanations for the restrictions they impose are consistent with our interpretation. At the end of the survey, we asked subjects why they constructed choice sets as they did. Subjects who withheld options typically cited paternalistic motives. Examples include, "*I wanted to force the future experiment participant to make the decision that is right for him*," "*A discount factor of 1/5 is very extreme. Other options offer a justifiable discount factor*," and "*I wanted to protect him from himself.*" Subjects who did not withhold options, in contrast, often described libertarian reasoning: "*The more options, the better*," and

[16] Data from the multiple decision lists are sometimes censored at the smallest and largest amounts specified in the lists. In column 3 of Table 3, 61 (7.04%) observations are censored above and 49 (5.66%) are censored below. In column 4, these figures are 25 (4.36%) and 99 (17.25%), respectively.

"*I will not dictate anyone's choice, even though I have a strong opinion.*"

Overall, we conclude that, for the most part, Choice Architects believe the restrictions they impose are beneficial, and therefore these interventions are consistent with paternalism. However, it does not necessarily follow that Choice Architects intervene *because* of this belief. Additional evidence, provided in subsequent sections, helps to bridge this gap. First, we show in Section 3.4 that interventions are sensitive to the provision of information that changes beliefs about the benefits of restrictions. Second, we provide evidence in Section 3.5 that causality does not flow from interventions to beliefs through a rationalization mechanism.

3.3 Do people intervene to correct for "present bias"?

So far, we have seen that many people intervene with benevolent intentions to prevent impatient choices that favor immediate payments over delayed rewards. These findings are consistent with the objectives that behavioral economists who believe decision makers suffer from *present bias*, defined as the tendency to favor immediate gratification excessively, commonly attribute to benevolent planners. Consequently, the next step in our analysis is to ask whether the alleviation of perceived present bias is in fact a central objective of the observed interventions.

As is well-known, introducing front-end delay (that is, delaying all payments by a fixed amount of time) induces people to choose more patiently for themselves (Frederick et al., 2002). Under the hypothesis that people suffer from present bias, this effect occurs because front-end delay removes the lure of immediacy that is responsible for excessively impatient choice. Thus, if people believe others suffer from present bias, front-end delay reduces the perceived benefits of paternalistic intervention. Indeed, under the particular positive and normative assumptions behavioral economists commonly invoke, front-end delay eliminates those benefits entirely.[17] Thus, if Choice Architects share the perspectives that behavioral economists often attribute to benevolent planners – if they indeed seek to correct for perceived present bias – the introduction of front-end delay should reduce or eliminate interventions. As we will see, our data reject this hypothesis in favor of the view that Choice Architects simply disapprove of impatient choices.

Focusing on the decisions Choice Architects make for themselves in Stage 2 of the laboratory session, we begin by replicating the usual finding that front-end delay increases patience. Column 1 of Table 4 presents a regression of the early payment the Choice Architect selects for herself on an indicator for front-end delay. The regression employs all of the decisions for menus 1 and 2 in the Main condition. Recall that each Choice Architect makes one of these two decisions, selected at random, with front-end delay. We include session, order, and menu fixed effects, and cluster standard errors by subject. As expected, the estimates show that the addition of front-end delay yields a statistically significant decrease in the selected early payment (€0.31, which corresponds to 5.6% of the greatest possible treatment effect).

Next, we ask whether Choice Architects anticipate Choosers' responses to front-end delay. Recall that,

[17] Most of the pertinent literature assumes that people have quasihyperbolic preferences, which take the following form: $U_t(x_t, x_{t+1}, ...) = u(x_t) + \beta \sum_{k=1}^{\infty} \delta^k u(x_{t+k})$, with $\beta, \delta < 1$. A common normative interpretation of this model is that δ captures true intertemporal preferences, while any deviation of β from unity constitutes a bias. In that case, a benevolent planner would try to maximize $\sum_{t=0}^{\infty} \delta^t u(x_t)$ (the long-run criterion). Bernheim and Taubinsky (2018) provide a critical evaluation of the foundations for this criterion.

Table 4: Effects of front-end delay and information about the Chooser.

	(1)	(2)	(3)	(4)	(5)
VARIABLES	€ Choice Architect takes early	Belief € Chooser takes early	Max. € Chooser takes early	Max. € Chooser takes early	Belief € Chooser takes early
Front-end delay	-0.312***	-0.228***	-0.104		
	(0.088)	(0.049)	(0.093)		
Chooser information					
Patient, happy				-0.454***	-1.480***
				(0.084)	(0.061)
Patient, unhappy				-0.049	-1.057***
				(0.072)	(0.054)
Impatient, unhappy				-0.312***	-0.223***
				(0.076)	(0.043)
Mean of dep. var.	0.965	1.929	4.548	3.246	1.521
	(0.098)	(0.062)	(0.082)	(0.058)	(0.045)
Observations	606	606	606	1,212	1,212
Number of subjects	303	303	303	303	303

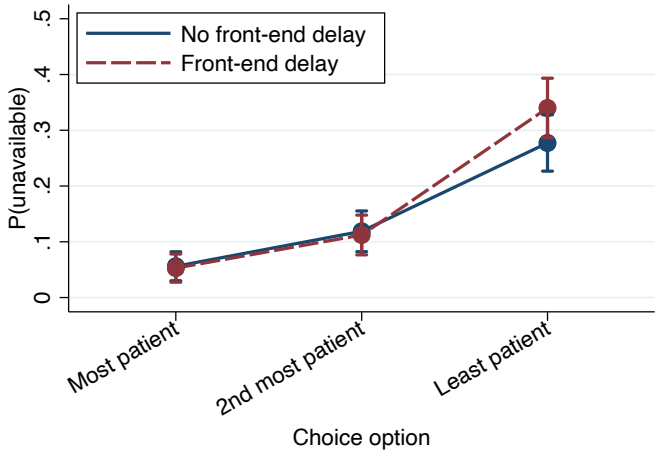
Notes: Method: OLS. Unit of observation: subject-round pair. Dependent variables: € *Choice Architect takes early* is the amount of Euros the Choice Architect chooses to receive early when choosing for herself in stage 2 of the laboratory component. *Belief Chooser takes early* is the Choice Architect's beliefs about the mean amount a Chooser will receive early if allowed to choose without restrictions. *Max. € Chooser takes early* is the maximal amount of Euros the Choice Architect permits the Chooser to receive early. Controls: *Front-end delay* is an indicator for whether all payments in a menu are delayed by a week. *Chooser information* is a set of dummies that indicate the statement a Chooser selects from Table 5 to describe himself. For columns 4 and 5, the omitted category is *impatient, happy*. Other controls include session, order, and (for columns 1-3) menu fixed effects. Samples: columns 1-3 use menus 1 and 2 from the Main condition; columns 4-5 use all rounds of the Chooser Information condition. Standard errors: clustered by subject. $^{*}p < 0.1$,$^{**}p < 0.05$,$^{***}p < 0.01$.

in Stage 2 of the laboratory component, we elicit Choice Architects' beliefs about the selections Choosers would make from all unrestricted menus. To generate the results in column 2 of Table 4, we regress the mean of the early payment for this elicited distribution on the same set of variables used in column 1. The results show that Choice Architects expect unrestricted Choosers to exhibit more patience with front-end delay, although the effect is slightly smaller than the impact on their own choices (€0.23).

How does front-end delay affect mandates? Figure 4 shows the frequency with which Choice Architects withhold each type of choice, averaged over all rounds in the Main condition that involve menus 1 and 2 (half of which include front-end delay). If anything, the interventions are slightly *more* common with front-end delay, not less. Column 3 of Table 4 makes the same point in a regression format. For this purpose, we define a Choice Architect's *mandate* as the largest early payment she allows the Chooser to receive. For example, if a Choice Architect offers the most patient and middle alternatives from menu 1 of Table 1, the mandate is €3.[18] Here we have regressed mandates for the same rounds on the same set of variables as in columns 1 and 2. The estimated effect of front-end delay is negative, meaning that mandates with front-end delay are more restrictive, not less, although the effect is not statistically significant. Further, Choice Architects

[18] All our results are robust with respect to alternative definitions of this variable, such as the smallest amount of money the Choice Architect forces the Chooser to receive with delay.

Figure 4: Effect of front-end delay.



Notes: Whiskers display 95%-confidence intervals with estimates of standard errors clustered at the subject level. Graph based on all decisions in the Main condition for menus 1 and 2.

are significantly *more* likely to think that their chosen restrictions are beneficial in the presence of front-end delay than without it, as Appendix B.7 shows. Thus, we find no evidence that Choice Architects intervene to counter present bias. Rather, they appear to regard patience as generally virtuous, and intervene to enforce it.

3.4 Are interventions sensitive to information about the affected party's inclinations and subjective experiences?

Having ruled out the possibility that our Choice Architects act on the types of objectives and perspectives that behavioral economists often attribute to benevolent planners, we consider a second, simpler hypothesis: paternalists make decisions for others based exclusively on their own conceptions of right and wrong without regard to the preferences or subjective experiences of the affected individuals. As we explain in this section, our data strongly reject this hypothesis. This finding leads naturally to the question we address in Section 4.2: In light of the fact that paternalists' beliefs about others' inclinations and subjective experiences affect their interventions, how do they form those beliefs?

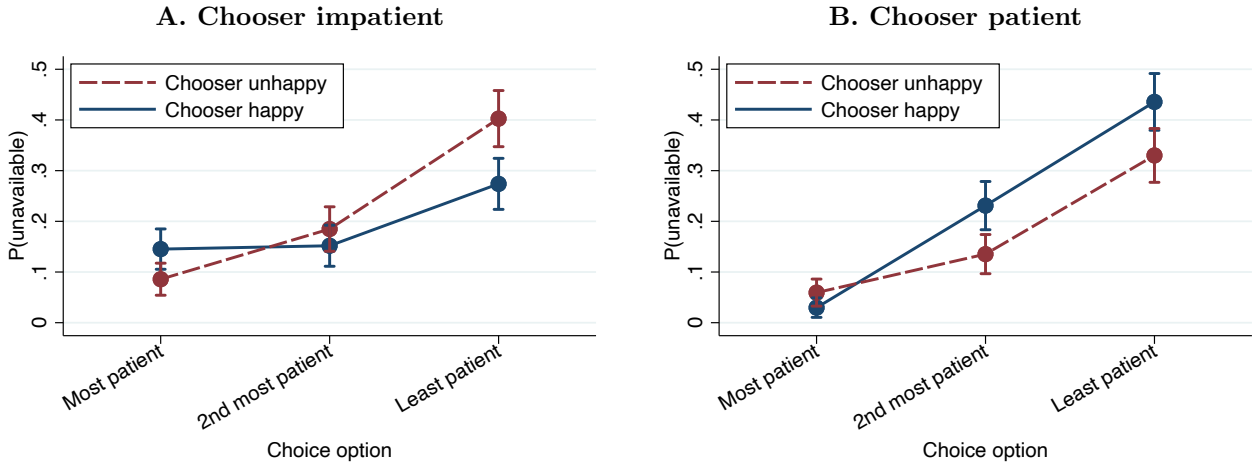
Table 5: Statements for the Chooser Information condition.

- I am a patient person. I am happy with this. (I often forego things in the present with regard to the future).
- I am an impatient person. I am happy with this. (I rarely forego things in the present with regard to the future).
- I am a patient person. I often regret my decisions. (Perhaps too often, I forego things in the present with regard to the future).
- I am an impatient person. I often regret my decisions. (Perhaps too rarely, I forego things in the present with regard to the future).

The analysis of this section employs data from the *Chooser Information* condition, which we mentioned briefly in Section 2.2. It resembles the Main condition, except that we provide Choice Architects with

information about their Choosers. Specifically, before making the decisions described in Section 2.4, each Chooser endorses one of the four statements listed in Table 5, which describe them as either patient or impatient, and as either generally happy or unhappy with their intertemporal choices. The condition consists of four rounds, one for each of the four statements. The Choice Architect knows her Chooser will fall into one of these categories, but does not know which one. Moreover, Choice Architects believe that a substantial fraction of Choosers falls into each of the four categories.[19] Consequently, they have reason to treat their decisions in each round as consequential. All four rounds involve menu 5 (see Table 1).[20]

Figure 5: Effect of information about the Chooser.



Notes: Whiskers indicate 95% confidence intervals with standard errors clustered by subject. Graph based on data from the Chooser Information condition, in which decisions involve menu 5.

Figure 5, which displays the frequency of withheld options for each Chooser statement, establishes that Choice Architects respond to this information. Panel A shows interventions affecting Choosers who describe themselves as impatient. Choice Architects impose more patience if an impatient Chooser claims to be unhappy with his impatience than if he claims to be happy with it. Thus, Choice Architects appear to select interventions that ostensibly help Choosers achieve their own goals. For Choosers who describe themselves as patient (Panel B), this pattern reverses. Choice Architects impose more patience if a patient Chooser claims to be happy with his patience (perhaps to prevent the Chooser from accidentally selecting impatient options) than if he claims to be unhappy with it.

Column 4 of Table 4 makes these same points in a regression format. The dependent variable is our measure of the Choice Architect's mandate, defined as in the Section 3.3. Using data from the Chooser Information condition, we regress mandates on indicators for each of the four Chooser self-descriptions (omitting one), as well as session and order fixed effects, clustering standard errors at the subject level. The results establish that Choice Architects' mandates differ significantly, both economically and statistically, across the four

[19] On average, Choice Architects believe that 33.1% of Choosers will classify themselves as patient and happy, 21.3% as impatient and happy, 16.5% as patient and unhappy, and 29.1% as impatient and unhappy. We elicited these beliefs either directly before or directly after eliciting beliefs about the distribution of choices by ten unrestricted subjects (randomized on the individual level).

[20] We ask subjects in the role of Choosers to select the statement that describes them best at the start of their session, before they receive any other information. If the Chooser Information condition is selected for implementation, the Chooser receives the choice set the Choice Architect constructed for the corresponding statement.

types of Choosers.[21]

Next we explore the mechanisms through which the information we provide affects interventions. The regression in column 5 of Table 4 shows how Choice Architects' beliefs about Choosers' unrestricted choices differ across the four self-descriptions. Recall that, in Stage 2 of the laboratory component, we elicit Choice Architects' beliefs about the selections Choosers would make from all unrestricted menus. Here we regress the mean of the early payment for this elicited distribution on indicators for the self-descriptions. The large and statistically significant differences between the four groups establish that the information changes Choice Architects' beliefs about the Choosers' inclinations. However, it is also apparent from the same regression that the information does not affect Choice Architects' interventions *only* by changing beliefs about Choosers' selections: conditional on knowing that the Chooser is unhappy, also knowing that he is patient rather than impatient drives beliefs and mandates in *opposite* directions (compare the pertinent coefficients in columns 4 and 5). The product of the differences in coefficients is negative and highly statistically significant ($p < 0.01$).[22]

3.5 Potential confounds

Paternalism is not the only possible motivation for intervention. To test for specific spurious motives, we supplemented our experiment with some additional conditions, which we briefly summarize in this section. Finding little evidence of spurious motives, in particular for interventions that Choice Architects regard as helpful to Choosers, we conclude that our results primarily reflect paternalism. To conserve space, details of these analyses appear in Appendix B.1.

The *Induced Chooser Preferences* condition, which we mentioned briefly in Section 2.2, addresses the concern that Choice Architects may intervene out of a general desire either to keep busy or to exercise active control over Choosers. By deploying Vernon Smith's induced preferences paradigm (Smith, 1976), it removes reasons to take issue with the Chooser's objectives while preserving the spurious motives mentioned above, as well as paternalistic motives unrelated to disapproval of preferences, such as the preemption of accidental errors.

In this condition, each choice option consist of a pair of immediate payments, one in Euros, the other in an experimental currency that converts to Euros at the rate $\hat{\delta}_i$, which is known only to Chooser i. Choice Architects merely know the population distribution of exchange rates. Formally, these decision problems resemble the intertemporal problems in the Main condition: one can think of the two payments as "immediate" and "delayed" (even though we pay both out right after the session) and interpret $\hat{\delta}_i$ as the Chooser's discount factor. The main difference is that, in this setting, the Choice Architect has no basis for

[21] Appendix B.7 shows how Choosers' self-descriptions relate to Choice Architects' beliefs about the welfare effects of chosen restrictions. Interventions continue to be consistent with benevolent motives. The likelihood with which Choice Architects see their interventions as helpful is highest when the Chooser is impatient and generally unhappy with his choices, second highest when the Chooser is patient and generally happy with his choices (possibly because Choice Architects remove options they think Choosers might select by accident), and lowest when the Chooser is either impatient and happy or patient and unhappy (apparently because Choice Architects are disinclined to enforce impatience).

[22] To perform this test, we estimate columns 4 and 5 jointly in a two-equation system OLS-regression with bootstrapped standard errors clustered on the subject level (1,000 samples).

disputing the normative validity of the Chooser's $\hat{\delta}_i$, because we have exogenously induced it. In contrast, the use of induced preferences does not remove motivations for intervention stemming from the desire to keep busy or to exercise active control over Choosers. There are two rounds in this condition using, respectively, analogs of menu 3 and menu 4 (see Table 1).

Choice Architects are substantially less likely to withhold options in the Induced Chooser Preferences condition than in the Main condition. Moreover, we do not see a pattern in the Induced Chooser Preferences condition analogous to our finding in the Main condition that Choice Architects become more inclined to limit impatient choices when impatience is more costly. Choice Architects are also far less likely to believe that interventions benefit the Chooser in the Induced Chooser Preferences condition. Specifically, Choice Architects both restrict choice and describe the restriction as helpful in 26.0% in the Main condition (excluding the rounds with front-end delay), compared to only 11.7% of the rounds in the Induced Chooser Preferences condition. We conclude that the patterns of primary interest in the Main condition likely reflect paternalism arising from Choice Architects' disapproval of impatience rather than spurious motivations.

A limitation of the Induced Chooser Preferences condition is that it cannot rule out a more nuanced hypothesis: Choice Architects may indulge their desire to exercise active control only if they can rationalize their actions as beneficial. A second purpose of the Exogenous Restrictions condition, part of which we discussed in Section 3.2, is to allow us to address this possibility.

To understand the logic of our test, consider a Choice Architect who would have constructed the two-option opportunity set had the default consisted of the three-option set. If we remove the least patient option exogenously, constructing the two-option opportunity set no longer involves an exercise of active control. To affect the outcome, the Choice Architect must now remove the middle option, which she can potentially rationalize as beneficial on the grounds that patience is virtuous. Thus, if the motives for interventions are spurious, the frequency with which Choice Architects withhold the middle option should rise. Accordingly, this condition includes a round in which the Choice Architect can remove options from a three-option choice set (menu 6), and a round in which we first remove the least patient option in that set exogenously. To account for unrelated boundary effects and noisy choosers, we include two otherwise identical versions of these rounds in which the default opportunity set consists of the most patient option by itself, so that the Choice Architect adds options instead of removing them.[23] These rounds disable the mechanism of concern, because a Choice Architect who enjoys exercising power, but who regards patience as virtuous, cannot rationalize the addition of impatient options. We find no evidence that Choice Architects are primarily motivated by a desire to exert control whenever they can rationalize active interventions as beneficial.

[23] Throughout this condition, when all three options are available, the Choice Architect can withhold either the least patient option, or the least patient and middle options; when we remove the least patient option exogenously, the Choice Architects can withhold the middle option. Thus, the Choice Architects' alternatives are more restricted than in the Main condition. In addition, as discussed in Section 3, when all three options are available, elicited beliefs about welfare effects involve comparisons between the three-option set and either the two-option set consisting of the most patient and middle options, or the one-option set consisting of the most patient option. When we remove the least patient option exogenously, these elicited beliefs involve comparisons between the two-option set and the set consisting of the most patient option.

4 Projective paternalism

We have shown that Choice Architects often withhold options for paternalistic reasons. But how do they decide whether particular options are good or bad for Choosers? By definition, paternalists are hesitant to rely on the judgments implicit in Choosers' decisions, and indeed may even question whether Choosers are aware of their own best interests. And yet we have also seen in Section 3.4 that paternalists are sensitive to Choosers' preferences and subjective experiences. While Choice Architects evidently find this information relevant, we require them to make decisions about interventions without knowing much about the Choosers, as is often the case in real-world applications. In this section, we seek to clarify how Choice Architects arrive at their evaluations in such situations.

4.1 Conceptual framework

To draw sharp inferences concerning the mechanisms governing Choice Architects' beliefs about Choosers, we interpret their behavior through the lens of a formal theory. Here, we develop a model of the Choice Architect's problem, precisely define various paternalistic mechanisms, and then identify their distinctive empirical signatures so we can distinguish among them.

4.1.1 Setting

There are two agents, a Choice Architect (she) and a Chooser (he). Potential consumption opportunities for the Chooser are indexed by a real number $c \in \mathbb{R}$. The Choice Architect can set a lower bound (*mandate*), $\underline{r} \in [-\infty, \infty)$, on the set of available alternatives, where we interpret $\underline{r} = -\infty$ as the absence of a mandate. She can also set an upper bound (*cap*), $\overline{r} \in (-\infty, \infty]$, where we interpret $\overline{r} = \infty$ as the absence of a cap. The Chooser then selects an option in $[\underline{r}, \overline{r}]$. A surrogate choice is one in which the Chooser sets these bounds subject to the restriction that $\underline{r} = \overline{r}$; we use s to denote this common value. For our intertemporal choice application, the index captures the balance between earlier and later payments, and the Choice Architect's selection of restrictions determines the most and least patient choices a Chooser may make.

The Choice Architect's beliefs about Choosers Suppose the Choice Architect believes the Chooser would select c from the unrestricted set $\mathbb{R}$, but that option u would be best for him. We refer to u as the Chooser's *ideal* choice, and to $m = u - c$ as the Chooser's *mistake* – in each case, *according to the Choice Architect's subjective judgment*, which may or may not coincide with (i) the Chooser's judgment, (ii) the Choice Architect's belief about the Chooser's judgment, or (iii) any external judgment.[24]

The Choice Architect has imperfect information about the Chooser, and her beliefs about the Chooser's ideal choices and mistakes are given by a cumulative distribution function $F(u, m)$. Next we describe some additional assumptions concerning these beliefs.

In many domains, it is reasonable to assume that perceived mistakes are unidirectional. Concerning

[24] For the intertemporal choice setting, the reader may be tempted to think of u and m in terms of quasi-hyperbolic discounting and the long-run criterion, but it is important to remember that our Choice Architects do not generally embrace that perspective (see Section 3.3). These variables reflect whatever normative judgments they actually apply.

intertemporal choice, there appears to be a widespread perception that people struggle to act sufficiently patiently rather than sufficiently impatiently. Indeed, as we saw in Section 3.1, the dominant pattern by far in our experiment is for subjects to impose patience rather than impatience. Accordingly, letting P denote the probability measure induced by F, we have:

Assumption 1 $P(m \geq 0) = 1$.[25]

This assumption involves a sign convention: the index of the ideal option is always at least as large as that of the chosen option. Thus, for our intertemporal choice setting, higher values of the index, c, denote greater patience, while $\underline{r}$ and $\overline{r}$ denote, respectively, the lowest and highest degrees of patience the Choice Architect permits.[26]

We also make some regularity assumptions concerning the Choice Architect's beliefs. To facilitate the statement of these assumptions, we decompose F into two components, as follows. We assume the Choice Architect believes the fraction γ of Choosers make normatively valid choices: m is identically equal to zero and the CDF of u is given by $Q(u)$. Furthermore, we assume the Choice Architect believes the fraction $1-\gamma$ of Choosers make normatively flawed choices: m is strictly positive and the CDF of (u, m) is given by $H(u, m)$. Thus, $F(u, m) = \gamma I_{m \geq 0} Q(u) + (1-\gamma) H(u, m)$, where $I_{m \geq 0}$ equals 1 if $m \geq 0$ and 0 otherwise.[27]

With this structure in place, we can state our technical assumptions:

Assumption 2 Q and H are atomless CDFs with well-defined densities, and with $\text{supp}\,(Q) = (-\infty, \infty)$ and $\text{supp}(H) = (-\infty, \infty) \times [0, \infty)$. Furthermore, the mean of u is finite for both Q and H, and the mean of m is finite for H.

Objective function The Choice Architect constructs the Chooser's opportunity set to maximize her perception of the Chooser's expected welfare. Specifically, the Choice Architect believes that if a Chooser's selection c differs from his ideal u, he sustains a welfare loss of $-l(c-u)$. We assume that $l(z)$ is differentiable with $l(0) = 0$, and that z and $l'(z)$ have opposite signs, so that $z = 0$ maximizes $l(z)$. We also assume for technical simplicity that $|l'|$ is bounded above, and that it is bounded away from zero outside any neighborhood of 0.[28]

We can write the Choice Architect's objective function as follows:

$$W(\underline{r}, \overline{r}) = \int l\big(\varphi_{u,m}(\underline{r}, \overline{r}) - u\big) dF(u, m), \tag{1}$$

where $\varphi_{u,m}(\underline{r}, \overline{r})$ denotes the Choice Architect's belief about the selection a Chooser of type (u, m) will make when choosing from the restricted set $[\underline{r}, \overline{r}]$. We assume the Choice Architect believes the Chooser will

[25] Our results continue to hold when Assumption 1 is violated as long as $P(m < 0)$ is not too large.

[26] In a setting involving exercise, where the concern is that mistakes generally involve inactivity, higher values of c would correspond to more frequent and vigorous workouts. In a setting involving nutrition, where the concern is that mistakes generally involve excessive quantities of unhealthy foods, higher value of c would correspond to healthier consumption.

[27] Studies in Behavioral Public Economics often employ the parallel assumption that the population includes a non-negligible group of "rational" consumers along with "behavioral" consumers (see Bernheim and Taubinsky, 2018). Here our assumption concerns the Choice Architect's beliefs about Choosers, rather than their actual characteristics.

[28] This assumption guarantees the existence of various integrals used in the proofs of our formal results. We can relax the assumption as long as the existence of these integrals is maintained.

select $c = u - m$ if that option is available, and the next closest available option otherwise. Accordingly, $\varphi_{u,m}(\underline{r}, \overline{r}) = \max\{\min\{c, \overline{r}\}, \underline{r}\}$.

It would be straightforward to add a term, $\kappa(\underline{r}, \overline{r})$, to the objective function for the purpose of capturing the Choice Architect's feelings (either positive or negative) about restricting the Choosers' options, other than those arising from anticipated effects on the Chooser's welfare. For simplicity we focus on the case of $\kappa \equiv 0$.[29]

Our results concern the comparative statics governing the optimal restrictions $[\underline{r}^*, \overline{r}^*]$, defined as the values that maximize expression (1), as well as the optimal surrogate choices s^*.

The Choice Architect's own choices Our empirical analysis focuses in large part on the relation between the choices Choice Architects make for themselves and the restrictions they impose on others. Accordingly, let c^A denote the selection the Choice Architect makes for herself when faced with the same decision problem as above. She believes that option u^A would be ideal. If $u^A \neq c^A$, she acknowledges that her actual choice c^A involves a mistake $m^A = u^A - c^A$.

Our next assumption, which concerns the distribution of types among Choice Architects, will eventually allow us to connect Choice Architects' own choices, c^A, to their mandates, $\underline{r}^*$. Formally, define $G_{m^A}(\cdot|c^A)$ as the marginal distribution of mistakes, and $G_{u^A}(\cdot|c^A)$ as the marginal distribution of ideals, among Choice Architects who choose c^A for themselves. We say that a CDF Γ for a variable x is increasing in a parameter θ if, for all $\theta' > \theta''$, $\Gamma(\cdot|\theta')$ first-order stochastically dominates $\Gamma(\cdot|\theta'')$.

Assumption 3 $G_{u^A}(\cdot|c^A)$ is increasing in c^A and $G_{m^A}(\cdot|c^A)$ is decreasing in c^A.

Intuitively, given the identity $c^A = u^A - m^A$, a higher value of c^A tends to indicate that u^A is higher and m^A is lower. Formally, the assumption rules out a strong positive correlation between ideals and mistakes among Choice Architects.[30]

4.1.2 The paternalist's decision

We begin with the fundamental trade-off between commitment and flexibility inherent in the paternalist's optimization problem. If the Choice Architect knew everything about the Chooser including his preferences, she would form a view as to which alternative is best for him and then eliminate any potential for mistakes by removing all other options. However, when the Choice Architect is not perfectly informed about the Chooser, any restriction may end up ruling out the best choice. As a result, optimal interventions often

[29] Our formulation of the Choice Architect's problem, together with the assumption $\kappa = 0$, rules out three types of motives that could affect interventions. First, we exclude the possibility that Choice Architects intervene due to a desire to exert control, which we would represent by assuming that $\kappa(\underline{r}, \overline{r}) > 0$ for $\underline{r} \neq -\infty$ and/or $\overline{r} \neq \infty$. This assumption is consistent with the experimental results in Section 3.5. Second, we abstract from pure libertarianism, which we might otherwise model by assuming that, for all restrictions, κ is negative and extremely large in magnitude for some fraction of the population. Assuming this concern for autonomy is uncorrelated with other aspects of preferences, libertarians and non-libertarians will behave identically when forced to make surrogate choices for others, consistent with our experimental results in Section 4.2 below. Third, in our model, the Choice Architect is concerned only with the Chooser's outcome, not with his choice process. For example, Choice Architects do not restrict Choosers' opportunity sets to save them the cognitive costs of making more complex choices, or to spare them the disutility associated with overcoming temptation (as in Thaler and Shefrin (1981) or implicitly in Gul and Pesendorfer (2001)). Importantly, we structure our experiment in a way that eliminates motives for intervention such as the cognitive costs or temptation experienced by Choosers.

[30] The assumption is satisfied when mistakes and ideals are independently distributed, when they are negatively correlated, and when they are not too strongly positively correlated. If u^A and m^A follow a bivariate Gaussian distribution with variances $\sigma^2_{u^A}$ and $\sigma^2_{m^A}$, respectively, and correlation ρ^A, then Assumption 3 is satisfied iff $\rho^A < \min\{\frac{\sigma_{u^A}}{\sigma_{m^A}}, \frac{\sigma_{m^A}}{\sigma_{u^A}}\}$.

leave the Chooser with a degree of autonomy. Moreover, under Assumption 1 (unidirectional mistakes), an upper bound on the choice of c imposes costs on Choosers with high values of u without preventing any mistakes. Accordingly, as shown in the following proposition, while the optimal intervention may involve a mandate ($\underline{r}^* > -\infty$), it never includes a cap (i.e., we always have $\overline{r}^* = \infty$).[31] The proposition also guarantees the existence of optimal mandates and surrogate choices. All proofs appear in Appendix A.

Proposition 1

(i) A Choice Architect does not impose a cap on the Chooser's options: $\overline{r}^* = +\infty$.

(ii) A Choice Architect may impose a mandate $\underline{r}^* > -\infty$ on the Chooser's options, but always leaves the Chooser with some flexibility: the set of optimizers is non-empty, closed, and bounded.[32]

(iii) A Choice Architect's optimal surrogate choice, s^*, is well-defined: The set of optimal surrogate choices is non-empty, closed, and bounded.

We cannot guarantee the uniqueness of the optimal mandate or surrogate choice without additional technical restrictions. In what follows, we assume that the Choice Architect breaks ties in favor of the Chooser's autonomy and thus selects the smallest maximizer (which, according to the proposition, is well-defined). The consistent application of other rules, such as selecting the largest maximizer, yields identical conclusions.

Varieties of paternalism We distinguish between various modes of paternalism, differentiated by the way in which the Choice Architect's beliefs about Choosers depend on her own type, (u^A, m^A). A *projective paternalist* forms her views about others by extrapolating from her understanding of herself. A Choice Architect exhibits *mistakes-projective paternalism* to the extent she thinks others are similar to her with respect to the mistakes they make. A Choice Architect exhibits *ideals-projective paternalism* to the extent she thinks others either have, or should have, ideals similar to her own. Naturally, a projective paternalist can be both mistakes-projective and ideals-projective. In contrast, a *conventional behavioral welfarist* acts as a benevolent social planner whose information about others may be incomplete, but who has accurate information about the underlying population distributions of key decision parameters, as envisioned in many standard behavioral welfare analyses (see, e.g., Bernheim and Taubinsky, 2018). She does not consider her own inclinations when intervening in others' choices.

To define these concepts precisely, we specialize to a semi-parametric representation of the Choice Architect's beliefs. Specifically, we assume that the distribution of beliefs over (u, m) is bivariate Gaussian, with parameters that depend on the Choice Architect's characteristics, as specified below. To maintain Assumption 1, we truncate m at zero:[33]

[31] For our intertemporal choice problem, given our sign convention, $\underline{r}^*$ is the minimum degree of patience the Choice Architect permits, and hence the maximum amount the Chooser can elect to receive early. Likewise, $\overline{r}^*$ is the maximum degree of patience the Choice Architect permits, and hence the minimum amount the Chooser can elect to receive early.

[32] The existence of residual flexibility hinges on our assumption that the underlying set of feasible options is unbounded. When the set of feasible options is bounded (as in our experiment), it may be optimal for the Choice Architect to eliminate all discretion.

[33] It is straightforward to generalize all of our results to a setting with two jointly Gaussian latent characteristics, z_u and z_m, such that u is a function of z_u, $u(z_u)$, and m is function of z_m, $m(z_m)$, where u and m are increasing in their arguments, in each case strictly so on some open interval. This approach allows us to accommodate our assumption that $P(m \geq 0) = 1$ in a variety of ways without ruling out natural distributional forms. For example, as an alternative to truncation, one could then assume that $m(z_m) = \alpha \exp(z_m)$.

Assumption 4 Let $(u, z_m) \sim \mathcal{N}(\mu_u, \mu_m, \sigma_u, \sigma_m, \rho)$. Then $(u, m) = \big(u, \max\{0, z_m\}\big)$.

This assumption yields a mass point at $m = 0$ that corresponds to the share γ of Choosers who the Choice Architect believes behave according to their ideals. Thus it implies specific functional forms for Q and H. We now turn to our formal definitions.

Definition

(i) Choice Architects practice *mistakes-projective paternalism* if μ_m is strictly increasing and μ_u is weakly increasing in m^A.

(ii) Choice Architects practice *ideals-projective* paternalism if μ_u is strictly increasing and μ_m is weakly increasing in u^A.

(iii) A Choice Architect is a *conventional behavioral welfarist* if μ_u and μ_m are the actual population means (and hence independent of the Choice Architect's type).

Each of the first two definitions involves a *direct effect* and a *compensatory effect*. The direct effects are intuitive: to the extent Choice Architects assume others either share or should share their ideals, we would expect μ_u to be strictly increasing in u^A (and similarly for mistakes). To understand the compensatory effects, imagine Choice Architects have sufficient information about others' behavior to correctly identify the population average of unrestricted choice (i.e., the mean of c, μ_c). Then, if an increase in u^A produces an increase in μ_u, it must also produce a fully compensating increase in μ_m. Likewise, with imperfect information about μ_c, one would expect a partially compensating increase in μ_m. Similar considerations apply to the inferences about others Choice Architects draw from their own mistakes.[34]

Testable implications Our main theoretical result uses Assumption 3 to show that each mode of paternalism has a distinctive empirical signature involving the sign of the correlation between, on the one hand, the selections Choice Architects make for themselves, and on the other hand, their mandates and surrogate choices.

Proposition 2

(i) Through the mistakes-projective mechanism, the distributions of the optimal mandate $\underline{r}^*$ and optimal surrogate choice s^* conditional on the Choice Architect's own selection, c^A, are decreasing in c^A (weakly in the case of surrogate choices).[35]

(ii) Through the ideals-projective mechanism, the distributions of the optimal mandate $\underline{r}^*$ and optimal surrogate choice s^* conditional on the Choice Architect's own selection, c^A, are increasing in c^A.

(iii) A conventional behavioral welfarist's optimal mandate $\underline{r}^*$ and optimal surrogate choice s^* are independent of her own choice c^A.

[34] Our notions of projective paternalism assume that the inferences a Choice Architect draws from her own characteristics pertain to the means, μ_u and μ_m, rather than to the variances or correlations, which we treat as invariant across Choice Architects. For the sake of tractability, we abstract from all forms of heterogeneity among Choice Architects other than in u^A and m^A.

[35] Recall that, when we say a distribution is increasing or decreasing in a parameter, we mean in the sense of first-order stochastic dominance.

The intuition for Proposition 2 is straightforward. Under Assumption 3, an increase in c^A generates an upward shift in the distributions of u^A and a downward shift in the distribution of m^A. Concretely, a consumer who acts more patiently probably aspires to greater patience and is less prone to mistakes involving excessive impatience. Through ideals-projective paternalism, the increase in u^A leads to an upward shift in the distribution of u (the main effect), and possibly an upward shift in the distribution of m (the compensatory effect). Both of these effects result in higher mandates – concretely, greater enforcement of patience. Through mistakes-projective paternalism, the decrease in m^A leads to a downward shift in the distribution of m (the main effect) and possibly a downward shift in the distribution of u (the compensatory effect). Both of these effects result in lower mandates – concretely, less enforcement of patience. As a result, greater patience when acting for oneself goes hand-in-hand with the imposition of greater patience to the extent Choice Architects are ideals-projective paternalists, and with the imposition of less patience to the extent they are mistakes-projective paternalists. For a conventional behavioral welfarist, own choices play no role. The intuition for surrogate choices is essentially the same, except that surrogate choice disables effects operating through m.

Accordingly, in the next subsection, we examine the empirical relationships between, on the one hand, the patience Choice Architects display when choosing for themselves and, on the other hand, their mandates and surrogate choices. We ask whether these relationships are consistent with the signature patterns implied by ideals-projective or mistakes-projective paternalism.

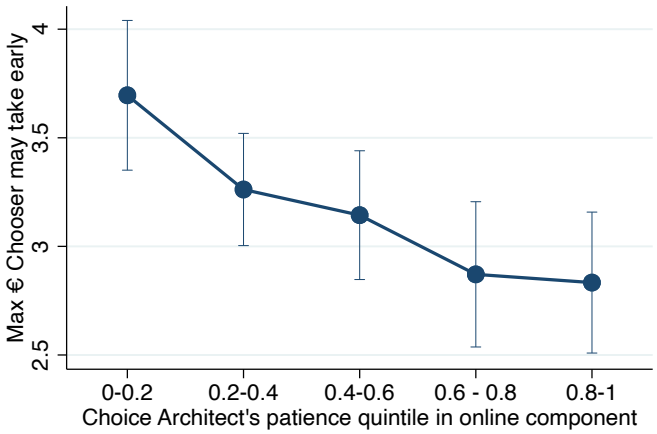
4.2 Distinguishing between variants of projective paternalism

The relationship between patience and mandates We have seen that mistakes-projection induces a *negative* relation between the options Choice Architects select for themselves and those they force on others, whereas ideals-projection induces a *positive* relation, and conventional behavioral welfarism implies the absence of a relation. To differentiate between these hypotheses, we study this relationship empirically. First we construct a measure of the degree of patience the Choice Architect displays when choosing for herself in the experiment's online component. Specifically, we calculate the percentile rank of the number of months she is willing to delay the receipt of the larger payment averaged over the six decision lists. Measuring patience in this way enables us to avoid assumptions about the structure of Choice Architects' intertemporal preferences. To avoid ambiguity, we focus on the 291 (of 303) Choice Architects who respected monotonicity in all of these lists.[36] Next we relate this measure of patience to Choice Architects' interventions in the Main condition, using the *Mandate* variable defined in Section 3.3 (i.e., the maximum amount the Choice Architect permits the Chooser to receive immediately).

Figure 6, which excludes libertarian subjects (38% of our sample), depicts our main result: those who have chosen more patiently for themselves in the online component of our experiment also impose significantly more patience on Choosers. The most patient non-libertarian Choice Architects' mandates are stricter than

[36] The 3.97% of subjects with multiple switches is low compared to other studies using multiple-decision lists, such as Holt and Laury (2002).

Figure 6: Ideals-projective paternalism.



Notes: Dependent variable: Maximum amount of money the Chooser is allowed to receive immediately as function of the Choice Architects' own patience. Sample: We exclude subjects classified as libertarian. Patience percentiles are unrelated to libertarianism, see Table 6. Whiskers indicate 95% confidence intervals.

those of the least patient Choice Architects by about €1. This difference is just under two-thirds of the standard deviation of mandates (€1.64) across this population, and one-fifth of the greatest possible effect (€5) in our experiment. According to Proposition 2 in Section 4.1.2, this finding is consistent with the signature pattern associated with ideals-projective paternalism.

Table 6: Relationships between Choice Architects' patience and measures of their mandates, surrogate choices, and welfare beliefs.

VARIABLES	(1) Mandate	(2) Mandate	(3) Libertarian	(4) Surrogate choice	(5) Surrogate choice	(6) Welfare belief
Non-libertarian subjects only	Yes			Yes		Yes
Patience %-ile	-1.044***	-0.614***	0.011	-1.589***	-1.655***	0.509***
	(0.270)	(0.197)	(0.091)	(0.242)	(0.183)	(0.130)
Mean of dep. var.	3.080	3.633	0.385	0.876	0.865	0.285
	(0.091)	(0.067)	(0.029)	(0.088)	(0.069)	(0.044)
Observations	537	873	291	518	837	537
Number of subj.	179	291	291	179	291	179

Notes: Method: OLS. Unit of observation: subject-round pairs for column 1, 2, 4, 5 and 6; subjects for column 3. Dependent variables: *Mandate* is the maximum amount the Choice Architect allows the Chooser receive immediately (as in Section 3.3). *Libertarian* is a dichotomous variable indicating that the Choice Architect never withholds options. *Surrogate choice* is the amount the Chooser receives immediately if the Choice Architect's surrogate choice is implemented. *Welfare belief* is the Choice Architect's belief about the welfare effect of the chosen restriction (as in Section 3.2). Controls: *Patience %-ile* is the Choice Architect's percentile rank according to the average number of months she is willing to delay the receipt of the larger payment in the online tasks. Other controls include session, order, and (with the exception of columns 3 and 6) menu fixed effects. Samples: All regressions limited to subjects who responded monotonically to all multiple-decision lists in the online component. Columns 1, 2, and 3 based on rounds from the Main condition without front-end delay. Columns 4 and 5 based on the corresponding surrogate choice rounds from Stage 2 of the laboratory component. The number of observations is smaller for surrogate choices because some of these choices were not recorded in the first two sessions. Column 6 based on the same two rounds of the Exogenous Restrictions condition used in Section 3.2. Columns 1, 4, and 6 restricted to non-libertarian subjects. Standard errors: clustered at the subject level. $^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$.

We formalize these observations by regressing Choice Architects' mandates in the Main condition on their patience percentiles. The unit of observation is a single decision in a single round. We control for session,

order, and menu fixed effects, and cluster standard errors by subject. For the regression in column 1 of Table 6, we exclude libertarians. As Figure 6 suggests, we find that the mandate tightens by €1.04 as patience rises from the bottom of the population distribution to the top, and the effect is highly statistically significant. Column 2 shows that the effect is smaller (€0.61) but still highly statistically significant once all subjects, including the libertarians, are included. The estimated effect is attenuated because libertarians never intervene, and because a Choice Architect's patience percentile does not predict whether she is libertarian (see column 3, which reports a subject-level regression of a libertarian indicator variable on the Choice Architect's patience percentile).

The relationship between patience and surrogate choices Next we ask whether Choice Architects differ in their judgments about what is good for Choosers, or merely in their propensities to intervene based on those judgments. To address this question, we examine Choice Architects' *surrogate choices*. Specifically, as mentioned in Section 2.2, in Stage 2 of the laboratory session we require Choice Architects to select a single item for the Chooser from each menu encountered previously in the experiment. Here we study each Choice Architect's surrogate choices for all of the Main-condition menus, excluding the one with front-end delay. Because these surrogate choices force Choice Architects to intervene, they directly reveal judgments about what is good for the Chooser. Column 4 shows that among non-libertarian subjects, the relation between Choice Architects' surrogate choices and their patience percentiles is even stronger than the relation between their mandates and their patience percentiles. Thus, the latter relationship reflects differences in judgments about what is good for the Chooser, and not merely differences in the propensity to intervene. Column 5 replicates column 4 using the entire sample. As the coefficient of interest changes only slightly, we infer that libertarians are similar to non-libertarians with respect to their judgments about others, except for their willingness to intervene. The positive relation between surrogate choices and own choices is part of the signature pattern associated with ideals-projective paternalism.

The relationship between patience and welfare beliefs Next we ask whether patience on the part of Choice Architects is related to their beliefs about the benefits of exogenously specified restrictions. For this purpose, we use data from the Exogenous Restrictions condition, much as in Section 3.2. Regressing our measure of beliefs about the welfare effects of restrictions on our measure of the Choice Architect's patience along with session and order fixed effects, we find that more patient Choice Architects are more likely to believe patience-promoting interventions are beneficial (see column 6 of Table 6).[37] Though not a formal implication of Proposition 2, this finding goes hand-in-hand with the tendency for more patient Choice Architects to impose greater patience, and it suggests that the key comparative static is indeed a consequence of paternalism.

The relationship between front-end delay and mandates In Section 3.3, we rejected the hypothesis that Choice Architects intervene to control perceived present bias by showing that the introduction of front-end delay increases the value Choice Architects believe Choosers derive from restrictions. According to

[37] Using our incentivized measure of negative compensating variation, the corresponding coefficient estimate is 0.191 (s.e. 0.094).

our estimates, it also increases the restrictiveness of interventions, and while that effect is not statistically significant, we certainly find no evidence that interventions decline. A natural explanation for this initially puzzling phenomenon now emerges: it is simply a consequence of ideals-projective paternalism. First recall that the introduction of front-end delay causes Choice Architects to behave more patiently when choosing for themselves. Perhaps because they are not fully aware of their inconsistencies (as evidence from Fedyk, 2017, implies), they behave as if their ideal level of patience, u^A, is higher with front-end delay than without (consistent with our Assumption 3). Projecting this greater level of ideal patience onto others, they impose more restrictive mandates. Any compensatory change in their beliefs about others' susceptibility to mistakes reinforces this tendency.

4.3 Are interventions misguided?

The fact that people engage in projective paternalism does not necessarily mean their interventions are objectively misguided. For example, an ideals-projective paternalist could have correct beliefs about the choices others would make if given the freedom to choose, but nevertheless decompose each of those choices into an ideal component and a mistaken component based on her own subjective values. To conclude that projective paternalists err when tightening or loosening constraints on others' choices, we must first demonstrate that their beliefs about others are systematically mistaken, and then confirm that those beliefs impact their interventions.

We begin by demonstrating that Choice Architects suffer from *false consensus bias* with respect to their' beliefs about others' choices. For the regression shown in column 1 of Table 7, the dependent variable is the Choice Architect's beliefs about the mean amount a Chooser will receive immediately if allowed to choose without restrictions from a given menu, calculated from the distribution elicited in the relevant round of Stage 2 during the laboratory session (as in Section 3.3). Focusing on the menus presented in the Main condition without front-end delay, we relate this variable to the Choice Architect's own patience percentile, defined as in Section 4.2. As with all other regressions in the table, we exclude libertarian subjects, control for session, order, and menu fixed effects, and present standard errors clustered at the subject level. Striking evidence of the false consensus effect emerges, in the sense that people tend to believe others will behave as they themselves behave. Compared to the most patient Choice Architects, the least patient ones think unrestricted Choosers would elect to receive an additional €1.30 immediately (26% of the maximum, €5, that Choosers can receive immediately).

Next we ask whether Choice Architects' skewed perceptions of others' inclinations infect the restrictions they impose. We have already seen in Section 4 that Choice Architects tailor their interventions based on information they receive about Choosers' propensities. As a general matter, one would therefore expect to find a strong relationship between their beliefs about others' unrestricted choices and their chosen restrictions. The hypothesized relationship is precisely what we observe in column 2 of Table 7. For this regression, the dependent variable is the Choice Architect's mandate, defined once again as the maximum amount she allows

the Chooser to receive immediately (see Section 3.3). Focusing on the menus presented in the Main condition without front-end delay and excluding libertarian subjects, we relate this variable to a measure of the Choice Architects' beliefs about the Chooser's impatience. To maintain comparability of the coefficient estimate to the effect of the Choice Architects' own patience, we use the percentile rank of the inverse of the mean amount she believes the Chooser will receive immediately if allowed to choose without restrictions. We also control for the Choice Architect's patience percentile. Because the Choice Architect's own preferences could drive their interventions directly while also incidentally impacting their beliefs about others through the false consensus effect, the omission of this control variable could potentially induce spurious correlation between the Choice Architect's mandate and her beliefs. The significant negative coefficient for the beliefs percentile implies, as expected, that Choice Architects impose more restrictive mandates when they believe Choosers are more patient. Indeed, Choice Architects' mandates are more strongly related to their percentiles ranks according to their beliefs about Choosers' unrestricted selections than to their percentiles ranks according to their own preferences.[38]

Table 7: Accuracy and relevance of beliefs.

VARIABLES	(1) Belief mean	(2) Mandate	(3) Believed relevance	(4) Actual relevance	(5) More restrictive than believed
Patience %-ile	-1.272***	-0.540*	0.011	0.041***	0.320***
	(0.167)	(0.318)	(0.029)	(0.013)	(0.064)
Belief %-ile		-1.136***			
		(0.365)			
Mean of dep. var.	1.467	3.100	0.095	0.056	0.209
	(0.062)	(0.089)	(0.010)	(0.004)	(0.023)
Observations	537	537	537	537	537
Number of subjects	179	179	179	179	179

Notes: Method: OLS. Unit of observation: subject-round pairs. Dependent variables: *Belief mean* is the Choice Architect's belief about the mean amount a Chooser will receive immediately if allowed to choose without restrictions (as in Section 3.3). *Mandate* is the maximum amount the Choice Architect allows the Chooser to receive immediately (as in Section 3.3). *Believed relevance* is the frequency with which the Choice Architect's mandate would bind according to her beliefs about Choosers' unrestricted selections. *Actual relevance* is the frequency with which the Choice Architect's restriction actually binds according to Choosers' unrestricted selections. *More restrictive than believed* is a dichotomous variable indicating whether the actual relevance is greater than the believed relevance. Controls: *Patience %-ile* is the Choice Architect's percentile rank according to the average number of months she is willing to delay the receipt of the larger payment in the online tasks. *Beliefs %-ile* is the Choice Architects' percentile rank in terms of her beliefs about the Chooser's patience, as measured by the inverse of the mean amount the Chooser will receive immediately if allowed to choose without restrictions. Additional controls include session, order, and menu fixed effects. Samples: Regressions limited to non-libertarian subjects who responded monotonically to all multiple-decision lists in the online component. Column 1 based on rounds from Stage 2 of the laboratory component that elicited Choice Architects' beliefs concerning Choosers' selections for the same menus as in the Main condition (excluding the one with front-end delay). Columns 2-6 based on rounds from the Main condition without front-end delay. Columns 4 and 6 also use data on Choice Architects' beliefs concerning Choosers' selections, gathered in Stage 2 of the laboratory component, in combination with the Choice Architect's chosen restriction to construct the dependent variable. Columns 5 and 6 also use data on unrestricted choices, gathered in Stage 2 of the laboratory component, in combination with the Choice Architects' chosen restrictions to construct the dependent variable. Standard errors: clustered at the subject level. $^{*}p < 0.1,^{**}p < 0.05,^{***}p < 0.01$.

So far, we have shown that false consensus bias infects Choice Architects' beliefs about others' choices,

[38] A possible objection to the regression in column 2 is that, because the two percentiles are strongly correlated, the regression may spuriously load their effects on the one that contains less measurement error. Appendix B.8 displays the results of a two-stage least-squares version of the same regression that addresses this concern. We find that the coefficient of the beliefs percentile more than doubles, and it remains highly statistically significant, while the coefficient of the patience percentile changes sign, declines slightly in magnitude, and is statistically insignificant.

and that those beliefs are closely related to their mandates. It follows that systematic bias also infects Choice Architects' decisions about interventions, which means these decisions are objectively misguided, even accepting the Choice Architects' aims. For a stark illustration of this point, see columns 3 through 5 of Table 7. For column 3, the dependent variable is the "believed relevance" of the restriction the Choice Architect imposes, defined as the fraction of Choosers for whom the Choice Architect thinks the mandate would bind. Specifically, for each menu encountered in the Main condition (excluding the round with front-end delay), we identify the least patient choice the Choice Architect permits, and then calculate the frequency with which the Choice Architect believes unrestricted Choosers would select less patient options according to the subjective distribution elicited in Stage 2 of the laboratory session. We relate this variable to the Choice Architect's patience percentile and the usual fixed effects. As indicated in the table, these subjects believe, on average, that they force 9.5% of Choosers to choose more patiently.

The regression reveals no systematic relation between these probabilities and Choice Architects' own patience: more patient Choice Architects do not intend to impose restrictions that enforce patience with higher frequency. And yet, that is precisely what they do. For column 4, the dependent variable is the "actual relevance" of the Choice Architect's restriction, defined as the fraction of Choosers for whom the restriction would actually bind. Specifically, for each menu encountered in the Main condition (excluding the round with front-end delay), we identify the least patient option the Choice Architect permits, and calculate the frequency with which unrestricted individuals choose less patiently than that.[39]

Because more patient Choice Architects impose more stringent mandates, an increase in the Choice Architect's patience percentile significantly increases the fraction of Choosers affected by the selected restriction. As a result, more patient Choice Architects are substantially more likely to underestimate the restrictiveness of their interventions. A simple tabulation reveals that 29.5% of Choice Architects in the most patient (non-libertarian) quartile underestimate the restrictiveness of their mandates, compared with only 9.0% of those in the least patient quartile. This conclusion also emerges from the regression in column 5, which employs a dichotomous dependent variable indicating whether the actual effect of the Choice Architect's restriction is greater than she believes. The estimated coefficient for the Choice Architect's patience percentile is highly statistically significant, which means that greater patience implies a significantly greater propensity to underestimate the restrictiveness of chosen mandates.[40]

Because Choice Architects' biased beliefs influence their interventions, it is natural to wonder whether informational policies could correct these errors and thereby improve the interventions, at least from the Choice Architects' perspectives. We address this question in Appendix B.4. Consistent with an existing literature that documents the resistance of the false consensus effect to informational interventions (Krueger

[39] For the actual unrestricted choice frequencies, we use the distribution of selections Choice Architects make for themselves from each of the pertinent menus in Stage 2 of the laboratory session (see Appendix B.9). Because there is only one Chooser for every four Choice Architects and each Chooser makes a single unrestricted choice, estimated frequencies based on Chooser data would be noisy. Significantly, we select Choice Architects and Choosers from the same population, and Choice Architects are aware of this fact.

[40] While we are mainly interested in the effect of patience on misperceptions of restrictiveness, it is worth noting that most of our Choice Architects actually overstate the restrictiveness of their chosen mandates. We interpret this finding with caution, however, due to the well-known tendency for elicited beliefs to be biased towards uniformity. Here, average beliefs about the distribution of choices across types of options (56%, 25%, and 19% for the most, second-most, and least patient options, respectively) are more uniform than the actual frequencies (78%, 14%, and 8%), and this pattern is responsible for the apparent prevalence of overstated restrictiveness.

and Clement, 1994; Engelmann and Strobel, 2012), we find that neither the Chooser-specific information provided in the Chooser Information condition, nor information about the distribution of unrestricted choices, attenuates the false consensus bias or ideals-projective paternalism.

5 Support for paternalistic policies

We conclude this investigation by showing that subjects' judgments about real-world paternalistic policy proposals relate meaningfully to the decisions they make as Choice Architects in our experiment. Additionally, we demonstrate that ideals-projective paternalism extends to the policy domain.

To this end, in Stage 3 of the experiment, we ask subjects to rate four policy proposals involving taxes on sugary drinks, alcohol, and tobacco, as well as restrictions on short-term, high-interest loans. Because our subjects live in Germany, we focus primarily on tax policies for Switzerland, which should purge the ratings of personal interests, at least in principle. We ask subjects to assume the tax policies would be budget-neutral, so responses do not reflect general attitudes about the size of government. For each policy, we elicit the extent to which the subject supports or opposes its implementation.[41] We also elicit beliefs about the impact of each policy on the welfare of the average citizen.[42]

After subjects express these judgments and answer additional non-incentivized questions, they provide information about themselves that relates to the impacted activities. These variables allow us to distinguish between ideals-projective paternalism and mistakes-projective paternalism in the domain of real-world paternalistic policy proposals. Specifically, we elicit subjects' body mass index,[43] their average alcohol consumption, their frequency of binge drinking (defined as the consumption of four or more units of alcohol for men, or five or more units for women, within a two-hour period),[44] their cigarette consumption, and their experience with short-term, high-interest loans. In addition, subjects provide information about their credit card debt in the online portion of the experiment (see Appendix D.1 for details).[45]

Relation of policy support to laboratory behavior For each policy proposal, we list the fractions of subjects who express a given level of support or opposition, as well as the fractions who provide each possible welfare assessment, in the top half of Table 8. Support outweighs opposition in each case. It is particularly pronounced for increased tobacco taxes and sugary drinks taxes. The distribution of beliefs about the welfare effects of the proposed policies mirrors the distribution of support for each policy.

To examine the relation between these judgments and the decisions subjects make as Choice Architects in

[41] Concerning tax policies, response options are that Switzerland should "definitely not introduce such a tax," "probably not introduce such a tax," "probably introduce such a tax," or "definitely introduce such a tax." Because small, short-term, high-interest loans are not available in Switzerland (possibly due to a lack of demand), questions regarding lending restrictions pertain to Germany. The response options are that the market for short-term loans in Germany should "be greatly restricted," "be somewhat restricted," "remain unchanged," "be somewhat liberalized," or "be greatly liberalized." We encode these options as 2, 1, 0, −1, and −2, respectively.

[42] The question about alcohol taxes concerns adolescents and young adults rather than average citizens, but is otherwise identical.

[43] Subjects can click a button to open a window that asks them to enter their height h in cm and weight w in kg. The window then displays their body mass index using the formula BMI $= \frac{w}{(h/100)^2}$.

[44] The experiment defines a unit of alcohol as 0.2 liters of beer, 0.1 liter of wine, or 1 shot of schnapps or liquor.

[45] Misreporting of these characteristics is a potential concern. For example, people tend to over-report height and under-report weight (Gorber et al., 2007). A strong correlation remains, however, between reported and measured BMI (Nawaz et al., 2001). Because our interest centers on the signs of correlations, underreporting of BMI does not qualitatively affect our conclusions. Similar statements apply to self-reported alcohol consumption (Sobell and Sobell, 1995) and self-reported cigarette smoking (West et al., 2007).

our experiment, we regress measures of their support for the policies on the average mandates they impose on Choosers. In the interests of obtaining easily interpretable coefficients, we use OLS, and defer ordered probit estimates to Appendix B.10. We encode strong opposition as -2, weak opposition as -1, weak support as 1, and strong support as 2.[46] We relate these measures of policy support to the Choice Architects' mandates (i.e., the maximum amount the Choice Architect allows the Chooser receive immediately, as in Section 3.3), averaged over rounds involving menus 3 and 4 in the Main condition.[47] Other controls include gender, age, field of study, high school GPA, monthly expenses, political orientation, and personal behaviors related to each of the policy domains listed above, as well as session and order fixed effects.[48] Panel B.1 in Table 8 displays the results. Focusing first on average support for the four paternalistic policies (column 1), we find that the coefficient of interest is negative and statistically significant ($p < 0.01$). Thus, subjects who impose tighter restrictions on Choosers in the laboratory also express greater support for paternalistic policies in the real world. We also perform separate regressions for each of the policy proposals and report the key coefficient estimate from each in columns 2 to 5. In each case, we find that Choice Architects who enforce more patience in the laboratory express greater support for actual paternalistic interventions. These relations are significant at the 5% level for alcohol and tobacco taxes and at the 10% level for the remaining two policies.

Next we examine the relation between beliefs about the welfare effects of actual paternalistic policies, and those concerning exogenous restrictions on choice in the laboratory. To measure the latter beliefs, we encode the Choice Architects' statements as to whether an exogenously specified restriction would help Choosers, leave them equally well off, or hurt them, as 1, 0, and -1, respectively (as in Section 3.2). We average this variable across the four rounds of the Exogenous Restrictions condition. For the regressions shown in Panel B.2 of Table 8, we include the same control variables as in Panel B.1. On average across the four policies (column 1), subjects who believe choice restrictions are beneficial in the laboratory also tend to believe that paternalistic interventions are beneficial in practice. The same pattern emerges for each of the sin taxes (columns 2-4). For restrictions on short-term lending, we do not find statistically significant effects. Because German universities do not charge tuition, personal debt may be less salient than alcohol, tobacco, and sugary drinks for our student subjects.

While we cannot exclude the possibility that non-paternalistic considerations such as the prevention of externalities factor into subjects' assessments of the four policies, Choice Architects' selections and statements in the laboratory session are not subject to such confounds. Consequently, it is reasonable to conclude from the robust relationships documented in Table 8 that attitudes toward these policies depend in significant part on the types of paternalistic inclinations we document in the laboratory.

[46] For lending restrictions, we code neutrality as 0.

[47] We confine attention to menus 3 and 4 so that we can take an average over the same set of menus for each Choice Architect without including rounds with front-end delay.

[48] Subjects could decline to provide pieces of information such as their body mass index. In that case, we set the value of the corresponding variable equal to its population mean, and we include dummy variables to indicate such replacements. Appendix B.11 displays a version of Table 8 in which subject-specific control variables are excluded from the regressions. If anything, the inclusion of control variables leads to an increase in the magnitude of the coefficient estimates.

Table 8: Experimental decisions and support for real-world paternalistic policies.

Policy	(1) All	(2) Increase alcohol tax	(3) Increase tobacco tax	(4) Introduce sugary drinks tax	(5) Tighten restrictions on short-term lending
		A. Summary statistics for dependent variables			
Distribution of support					
strongly opposed (-2)	-	0.154	0.055	0.087	0.060
weakly opposed (-1)	-	0.305	0.102	0.196	0.191
neutral (0)	-	-	-	-	0.342
weakly in favor (1)	-	0.328	0.280	0.362	0.278
strongly in favor (2)	-	0.213	0.563	0.355	0.129
		B. Relation between laboratory choice and policy attitudes			
		B.1 Dependent variable: support for policy proposal			
Average mandate imposed on Chooser	-0.113***	-0.129**	-0.126**	-0.113*	-0.086*
(max. € paid immediately)	(0.041)	(0.066)	(0.051)	(0.066)	(0.052)
Observations	403	403	403	403	403
		B.2 Dependent variable: beliefs about welfare effect of policy proposal			
Average belief exog. smaller choice	0.185***	0.220**	0.235***	0.210***	0.075
set better for Chooser	(0.054)	(0.086)	(0.084)	(0.076)	(0.075)
Observations	403	403	403	403	403
		C. Projective paternalism with actual policies			
Alcohol consumption					
Alcohol units / week		-0.050			
		(0.034)			
log(days binge drinking / year)		-0.165***			
		(0.055)			
Tobacco consumption					
Smoker yes / no			-0.881***		
			(0.284)		
Cigarettes / day			-0.046		
			(0.039)		
Body Mass Index				-0.062**	
				(0.028)	
Debt					
Credit card debt (in €1,000)					-0.573***
					(0.114)
Other short-term debt yes / no					0.101
					(0.388)
Controls	-	Yes	Yes	Yes	Yes
Observations	-	403	403	398	351

Notes: Method: OLS. Unit of observation: subject. Dependent variables: *Support for policy proposal* is the support expressed for various policies, coded -2 (strong opposition) to 2 (strong support), averaged over policies for "All." A "neutral" response was possible only for the question about short-term lending. Because we asked subjects about *loosening* restrictions on short-term lending, we reverse-coded these responses for easier comparability (so that higher values correspond to greater support for tightening restrictions). *Beliefs about welfare effects* is the belief about the welfare effect of various policies, coded -2 (significantly worse off) to 2 (significantly better off), averaged over policies for "All." Controls: *Average mandate imposed on Chooser* is the maximum amount the Choice Architect allows the Chooser receive immediately (as in Section 3.3), averaged over rounds involving menus 3 and 4 in the Main condition. *Average belief exog. smaller choice set better for Chooser* (as in Section 3.2) encodes whether the Choice Architect considers the smaller opportunity set better, equally good, or worse for the Chooser than the unrestricted set, coded as 1, 0, and -1, respectively, averaged across the four rounds of the Exogenous Restriction Condition. All regressions control for gender, age, self-reported political attitudes, log monthly expenses, high school GPA, university faculty at which the subject's major field of study is offered, as well as for session and order fixed effects. Regressions in Panel B in addition control for weekly alcohol consumption, log days of binge drinking per year (defined as the consumption of at least 4 units of alcohol for females, 5 for males, within a period of two hours (National Institutes on Alcohol Abuse and Alcoholism, 2018)), smoking status, number of cigarettes smoked per day, body mass index, credit card debt, and for having taken a short term loan. For variables measured with interval precision, we use midpoints for analysis. For control variables that subjects chose not to disclose, the regressions in Panel B impute population means, and include indicators for whether a variable's value was missing. Samples: includes 303 subjects who participated in the main experiment, plus 100 subjects who participated in the Choice Distribution Information Condition (see Appendix B.4). Regressions in Panel C exclude subjects who chose not to disclose the personal characteristic of interest. Standard errors: clustered by subject. $^{*}p < 0.1$,$^{**}p < 0.05$,$^{***}p < 0.01$.

Projective paternalism with actual policies Next we investigate the relationships between respondents' policy judgments and their own characteristics, with the objective of distinguishing between mistakes-projective and ideals-projective paternalism. Focusing on the example of alcohol taxes, mistakes-projective paternalism should yield a positive relationship between alcohol consumption and support for alcohol taxation in politically unrelated jurisdictions, where the impact will fall on others. Intuitively, heavier drinkers are more likely to be problem drinkers. To the extent they project their problem drinking on others, they will conclude that others would benefit from measures that limit alcohol consumption. Ideals-projective paternalism has the opposite implication. Intuitively, light drinkers are less likely to enjoy consuming alcohol. To the extent they project their lack of enjoyment on others, they will conclude that the costs of limitating alcohol consumption are relatively low. Therefore, we should expect to observe a negative relationship between alcohol consumption and support for alcohol taxes.

To test these competing implications, we estimate OLS regressions relating subjects' support for each policy to pertinent personal behaviors and characteristics.[49] We include all subjects from all treatments, and include the same set of additional control variables as in the regressions of Table 8.

Panel C displays the results. Column 2 shows that subjects who binge drink more frequently are significantly *less* likely to express support for alcohol taxes on others, exactly as ideals-projective paternalism predicts. Because binge drinking is a good proxy for problem drinking, this finding strongly contradicts the hypothesis of mistakes-projective paternalism. Conditional on the level of binge drinking, greater weekly alcohol consumption also appears to reduce support for alcohol taxes on others, but the effect is not statistically significant.[50] Similarly, column 3 shows that German subjects are significantly more likely to express support for tobacco taxes in Switzerland if they do not smoke themselves. In column 4, we find that people with lower BMIs express stronger support for sugary drinks taxes on others, again consistent with ideals-projective paternalism. Similarly, subjects are less likely to support restricting other people's access to the market for short-term, high-interest lending when they have larger credit card balances (column 5).[51] While results such as those of Allcott et al. (2019) suggest that the relationships between policy support and subject characteristics may partially be driven by differences in domain-specific knowledge such as health literacy, this factor cannot plausibly explain the strong relationship between laboratory choice and policy attitudes documented in Panel B of the table. Overall, the data on judgments about real-world paternalistic policies therefore manifests the signature pattern associated with ideals-projective paternalism.

In principle, the observed relationship between one's own behavior and support for paternalistic policies may also be due to systematic variation in beliefs about the effectiveness of a given policy. Appendix C reports a vignette experiment with US subjects in which we control for variations in beliefs about efficacy. In cases where the pattern of support for paternalistic policies points to ideals-projective paternalism, controlling for

[49] Appendix B.10 presents corresponding estimates obtained from ordered probit models.

[50] The coefficient on weekly consumption becomes statistically significant (and remains negative) when we remove the control for binge drinking from the regression.

[51] Only 12 out of 403 subjects report ever having taken a short-term loan other than through their credit cards.

beliefs about efficacy leaves that pattern qualitatively unchanged.[52]

6 Conclusion

This paper examines when, why, and how people intervene in others' choices. In a setting involving intertemporal tradeoffs, Choice Architects frequently remove options that are attractive to impatient decision makers. They believe their interventions benefit the Chooser, and are thus acting paternalistically. We examine and reject two simple hypotheses about their motives. First, we rule out the possibility that Choice Architects intervene with the objective of controlling excessive impatience arising from perceived present bias, a motive practitioners of behavioral welfare economics often attribute to benevolent planners. Second, we rule out the possibility that Choice Architects intervene simply to impose their own judgments, without regard to the inclinations or subjective experiences of the affected individuals.

How do Choice Architects judge what is good for others? Ideals-projective paternalism emerges from our analysis as the key mechanism. An ideals-projective paternalist acts *as if* she believes others share, or ought to share, the ideals to which she aspires for herself. We show that ideals-projective paternalism is related to the false consensus effect (an objective fallacy). As a result, even though more patient Choice Architects do not intend to force a larger fraction of Choosers to choose more patiently, this is precisely what they do. We also find that ideals-projective paternalism extends to assessments of real-world paternalistic policies, which strongly correlate with behavior in the laboratory.

Throughout, we have remained agnostic about the effects of Choice Architects' interventions on Choosers' welfare. Finding an objective basis for making such assessments is challenging. For example, from a libertarian perspective, any intervention is welfare-reducing. Alternatively, if one believes that, given the high cost of impatience in our experiment, the most patient option always dominates the other alternatives, then removing the least patient option, or both the least patient and middle options, is weakly welfare-enhancing, because it helps Choosers avoid accidental errors. Existing evidence suggests, however, that people have a positive willingness to pay for autonomy (Fehr et al., 2013; Bartling et al., 2014; Owens et al., 2014; Lübbecke and Schnedler, 2018; Ackfeld and Ockenfels, 2020).

There are many questions we hope future research will clarify. For example, how much of their own resources are subjects willing to give up in order to impose paternalistic restrictions on others, and do these amounts differ for mistakes-projective and ideals-projective paternalists? We also hope to extend the empirical study of paternalistic decision making to subject pools consisting of "professional paternalists" such as medical doctors and policy makers. In contexts where objective benchmarks are available, existing evidence suggests that nearly everyone exhibits behavioral biases (Stango and Zinman, 2019), including elected politicians (Sheffer et al., 2018).

[52] That survey asks about alcohol taxes, retirement savings mandates, restrictions on short-term, high-interest lending and sugary drinks taxes. We observe statistically significant indications of ideals-projective paternalism for the first two policies. The evidence does not support mistakes-projective paternalism for any policy.

References

Ackfeld, Viola and Axel Ockenfels, "Protecting Autonomy versus Intervening to Promote Prosocial Behavior," *unpublished*, 2020.

Allcott, Hunt, Benjamin B Lockwood, and Dmitry Taubinsky, "Regressive sin taxes, with an application to the optimal soda tax," *The Quarterly Journal of Economics*, 2019, *134* (3), 1557–1626.

Altmann, Steffen, Armin Falk, and Andreas Grunewald, "Incentives and information as driving forces of default effects," *unpublished*, 2017.

Ambuehl, Sandro, "An Offer You Can't Refuse? Incentives Change How We Think," *CESifo working paper 6296*, 2017.

_ **and Axel Ockenfels**, "The Ethics of Incentivizing the Uninformed. A Vignette Study," *The American Economic Review Papers and Proceedings*, 2017, *107* (5), 91–95.

_ **, Muriel Niederle, and Alvin E. Roth**, "More Money, More Problems? Can High Pay be Coercive and Repugnant?," *American Economic Review, Papers & Proceedings*, 2015, *105* (5), 357–60.

Bartling, Björn, Alexander Cappelen, Henning Hermes, Marit Skivenes, and Bertil Tungodden, "Free to Fail? Paternalistic Preferences in the U.S.," *unpublished*, 2020.

Bartling, Björn, Ernst Fehr, and Holger Herz, "The intrinsic value of decision rights," *Econometrica*, 2014, *82* (6), 2005–2039.

Basu, Kaushik, "The economics and law of sexual harassment in the workplace," *The Journal of Economic Perspectives*, 2003, *17* (3), 141–157.

_ , "Coercion, Contract and the Limits of the Market," *Social Choice and Welfare*, 2007, *29* (4), 559–579.

Benartzi, Shlomo, John Beshears, Katherine L Milkman, Cass R Sunstein, Richard H Thaler, Maya Shankar, Will Tucker-Ray, William J Congdon, and Steven Galing, "Should governments invest more in nudging?," *Psychological Science*, 2017, *28* (8), 1041–1055.

Bennyhoff, Donald G and Francis M Kinniry Jr, "Advisor's Alpha," *Vanguard research paper (December). https://advisors. vanguard. com/iwe/pdf/ICRAA. pdf*, 2011.

Bernheim, B. Douglas, "The Good, the Bad, and the Ugly: A Unified Approach to Behavioral Welfare Economics," *Journal of Benefit-Cost Analysis*, 2016, *7* (1), 12–68.

_ **and Antonio Rangel**, "Beyond Revealed Preference: Choice-Theoretic Foundations for Behavioral Welfare Economics," *The Quarterly Journal of Economics*, 2009, *124* (1), 51–104.

Bernheim, B Douglas and Dmitry Taubinsky, "Behavioral public economics," in B. Douglas Bernheim, Stefano DellaVigna, and David Laibson, eds., *Handbook of Behavioral Economics: Applications and Foundations 1*, Vol. 1, Elsevier, 2018, pp. 381–516.

Besley, Timothy, "A simple model for merit good arguments," *Journal of Public Economics*, 1988, *35* (3), 371–383.

Bisin, Alberto, Alessandro Lizzeri, and Leeat Yariv, "Government policy with time inconsistent voters," *American Economic Review*, 2015, *105* (6), 1711–37.

Boven, Leaf Van, David Dunning, and George Loewenstein, "Egocentric empathy gaps between owners and buyers: misperceptions of the endowment effect.," *Journal of Personality and Social Psychology*, 2000, *79* (1), 66.

Carlin, Bruce Ian, Simon Gervais, and Gustavo Manso, "Libertarian paternalism, information production, and financial decision making," *The Review of Financial Studies*, 2013, *26* (9), 2204–2228.

Clemens, Michael A, "Testing for repugnance in economic transactions: Evidence from guest work in the Gulf," *The Journal of Legal Studies*, 2018, *47* (S1), S5–S44.

Currie, Janet and Firouz Gahvari, "Transfers in cash and in-kind: Theory meets the data," *Journal of Economic Literature*, 2008, *46* (2), 333–83.

Delavande, Adeline, Xavier Giné, and David McKenzie, "Measuring subjective expectations in developing countries: A critical review and new evidence," *Journal of Development Economics*, 2011, *94* (2), 151–163.

DellaVigna, Stefano and Ulrike Malmendier, "Contract design and self-control: Theory and evidence," *The Quarterly Journal of Economics*, 2004, *119* (2), 353–402.

Dworkin, Gerald, *Paternalism*, Wadsworth Publishing Company, 1971.

_ , "Paternalism," *the Monist*, 1972, pp. 64–84.

Elias, Julio J, Nicola Lacetera, and Mario Macis, "Markets and Morals: An Experimental Survey Study," *PloS one*, 2015, *10* (6).

Elias, Julio, Nicola Lacetera, and Mario Macis, "Sacred Values? The Effect of Information on Attitudes Toward Payments for Human Organs," *American Economic Review, Papers & Proceedings*, 2015, *105* (5), 361–365.

_ , _ , and _ , "Paying for kidneys? A randomized survey and choice experiment," *American Economic Review*, 2019, *109* (8), 2855–88.

Engelmann, Dirk and Martin Strobel, "Deconstruction and reconstruction of an anomaly," *Games and Economic Behavior*, 2012, *76* (2), 678–689.

Exley, Christine L and Judd B Kessler, "Equity Concerns are Narrowly Framed: Why Money Cannot Buy Time," *unpublished*, 2017.

Falk, Armin, Anke Becker, Thomas Dohmen, David Huffman, and Uwe Sunde, "The preference survey module: A validated instrument for measuring risk, time, and social preferences," 2016.

Faravelli, Marco, Priscilla Man, and Randall Walsh, "Mandate and paternalism: a theory of large elections," *Games and Economic Behavior*, 2015, *93*, 1–23.

Fedyk, Anastassia, "Asymmetric Naiveté: Beliefs about Self-Control," *unpublished*, 2017.

Fehr, Ernst, Holger Herz, and Tom Wilkening, "The lure of authority: Motivation and incentive effects of power," *The American Economic Review*, 2013, *103* (4), 1325–1359.

Foerster, Stephen, Juhani T Linnainmaa, Brian T Melzer, and Alessandro Previtero, "Retail financial advice: does one size fit all?," *The Journal of Finance*, 2017, *72* (4), 1441–1482.

Frederick, Shane, George Loewenstein, and Ted o'Donoghue, "Time discounting and time preference: A critical review," *Journal of Economic Literature*, 2002, *40* (2), 351–401.

Gangadharan, Lata, Philip J Grossman, and Kristy Jones, "Donor types and paternalism," *unpublished*, 2015.

Gigerenzer, Gerd, "Moral intuition = fast and frugal heuristics?," in "The cognitive science of morality: intuition and diversity," MIT Press, 2008, pp. 1–26.

Glaeser, Edward L, "Paternalism and psychology," *NBER working paper 11789*, 2005.

Gorber, S. Connor, M. Tremblay, D. Moher, and B. Gorber, "A comparison of direct vs. self-report measures for assessing height, weight and body mass index: a systematic review," *Obesity Reviews*, 2007, *8* (4), 307–326.

Gruber, Jonathan and Botond Kőszegi, "Is addiction "rational"? Theory and evidence," *The Quarterly Journal of Economics*, 2001, *116* (4), 1261–1303.

_ and _ , "Tax incidence when individuals are time-inconsistent: the case of cigarette excise taxes," *Journal of Public Economics*, 2004, *88* (9-10), 1959–1987.

Gul, Faruk and Wolfgang Pesendorfer, "Temptation and self-control," *Econometrica*, 2001, pp. 1403–1435.

Holt, Charles A. and Susan K. Laury, "Risk Aversion and Incentive Effects," *American Economic Review*, 2002, *92(5)*, 1644 – 1655.

Ifcher, John and Homa Zarghamee, "Behavioral Economic Phenomena in Decision-Making for Others," *The Journal of Economic Psychology*, 2020.

Jacobsson, Fredric, Magnus Johannesson, and Lars Borgquist, "Is altruism paternalistic?," *The Economic Journal*, 2007, *117* (520), 761–781.

Kahneman, Daniel, *Thinking Fast and Slow*, New York: Macmillan, 2011.

Kataria, Mitesh, M Vittoria Levati, and Matthias Uhl, "Paternalism with hindsight: do protégés react consequentialistically to paternalism?," *Social Choice and Welfare*, 2014, *43* (3), 731–746.

Krawczyk, Michal and Lukasz P Wozny, "An experiment on temptation and attitude towards paternalism," *unpublished*, 2017.

Krueger, Joachim and Russell W Clement, "The truly false consensus effect: An ineradicable and egocentric bias in social perception.," *Journal of personality and social psychology*, 1994, *67* (4), 596.

Laibson, David, "Private Paternalism, the Commitment Puzzle, and Model-Free Equilibrium," *American Economic Review, Papers & Proceedings*, 2018, *108*, 1–21.

Leider, Stephen and Alvin E Roth, "Kidneys for sale: Who disapproves, and why?," *American Journal of Transplantation*, 2010, *10* (5), 1221–1227.

Linnainmaa, Juhani T, Brian Melzer, and Alessandro Previtero, "The misguided beliefs of financial advisors," *Journal of Finance*, forthcoming.

Locke, John, *Two Treatises of Government*, A. Millar et al., London, 1764.

Loewenstein, George and Emily C Haisley, *The economist as therapist: Methodological ramifications of 'light' paternalism*, Vol. 1 of *Handbook of Economic Methodologies*, Oxford University Press, 2007.

_ , Ted O'Donogue, and Matthew Rabin, "Projection Bias in Predicting Future Utility," *Quarterly Journal of Economics*, 2003, *118* (4), 1209–1248.

Lübbecke, Silvia and Wendelin Schnedler, "Don't patronize me! An Experiment on Rejecting Paternalistic Help," *unpublished*, 2018.

Lusk, Jayson L, Stephan Marette, and F Bailey Norwood, "The paternalist meets his match," *Applied Economic Perspectives and Policy*, 2013, *36* (1), 61–108.

Madarasz, Kristof, "Information projection: Model and applications," *The Review of Economic Studies*, 2012, *79* (3), 961–985.

Mill, John Stuart, *On liberty*, Longmans, Green, Reader, and Dyer, 1869.

Moffitt, Robert, "Welfare work requirements with paternalistic government preferences," *The Economic Journal*, 2006, *116* (515), F441–F458.

Moffitt, Robert A, "The negative income tax and the evolution of US welfare policy," *Journal of Economic Perspectives*, 2003, *17* (3), 119–140.

Mulligan, Casey B and Tomas J Philipson, "Merit motives and government intervention: Public finance in reverse," *National Bureau of Economic Research Working Paper*, 2000, *7698*.

National Institutes on Alcohol Abuse and Alcoholism, "Recommended Alcohol Questions," https://www.niaaa.nih.gov/research/guidelines-and-resources/recommended-alcohol-questions July 2018.

Nawaz, Haq, Wendy Chan, Mustapha Abdulrahman, David Larson, and David L Katz, "Self-reported weight and height: implications for obesity research," *American Journal of Preventive Medicine*, 2001, *20* (4), 294–298.

O'Donoghue, Ted and Matthew Rabin, "Doing it now or later," *American Economic Review*, 1999, *89* (1), 103–124.

_ and _ , "Optimal sin taxes," *Journal of Public Economics*, 2006, *90* (10-11), 1825–1849.

Owens, David, Zachary Grossman, and Ryan Fackler, "The control premium: A preference for payoff autonomy," *American Economic Journal: Microeconomics*, 2014, *6* (4), 138–61.

Rizzo, Mario J and Glen Whitman, *Escaping Paternalism: Rationality, Behavioral Economics, and Public Policy*, Cambridge University Press, 2019.

Ross, Lee, David Greene, and Pamela House, "The "false consensus effect": An egocentric bias in social perception and attribution processes," *Journal of Experimental Social Psychology*, 1977, *13* (3), 279–301.

Roth, Alvin E, "Repugnance as a Constraint on Markets," *Journal of Economic Perspectives*, 2007, *21* (3), 37–58.

Saint-Paul, Gilles, "Policy consequences of limited cognitive ability," *Spanish Economic Review*, 2004, *6* (2), 97–105.

Schroeder, Juliana, Adam Waytz, and Nicholas Epley, "Endorsing Help For Others That You Oppose For Yourself: Mind Perception Alters the Perceived Effectiveness of Paternalism.," *Journal of Experimental Psychology: General*, 2017.

Sheffer, Lior, Peter John Loewen, Stuart Soroka, Stefaan Walgrave, and Tamir Sheafer, "Non-representative representatives: an experimental study of the decision making of elected politicians," *American Political Science Review*, 2018, *112* (2), 302–321.

Smith, Vernon L, "Experimental economics: Induced value theory," *The American Economic Review*, 1976, pp. 274–279.

Sobell, Linda C and Mark B Sobell, "Alcohol consumption measures," *Assessing alcohol problems: A guide for clinicians and researchers*, 1995, *2*, 75–99.

Stango, Victor and Jonathan Zinman, "We are all behavioral, more or less: Measuring and using consumer-level behavioral sufficient statistics," *NBER working paper*, 2019, *25540*.

Sunstein, Cass R, "Moral heuristics," *Behavioral and Brain Sciences*, 2005, *28* (4), 531–541.

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

_ **and Hersh M Shefrin**, "An economic theory of self-control," *Journal of Political Economy*, 1981, *89* (2), 392–406.

Uhl, Matthias, "Do self-committers mind commitment by others? An experiment on weak paternalism," *Rationality, Markets and Morals*, 2011, *2*, 13–34.

West, Robert, Witold Zatonski, Krzysztof Przewozniak, and Martin J Jarvis, "Can we trust national smoking prevalence figures? Discrepancies between biochemically assessed and self-reported smoking rates in three countries," *Cancer Epidemiology and Prevention Biomarkers*, 2007, *16* (4), 820–822.

Zamir, Eyal, "The efficiency of paternalism," *Virginia Law Review*, 1998, pp. 229–286.

ONLINE APPENDIX

What Motivates Paternalism? An Experimental Study

Sandro Ambuehl, B. Douglas Bernheim, Axel Ockenfels

Table of Contents

A Proofs

A.1 Proof of Proposition 1

Mandates and caps We let q and h denote the density functions of Q and H, respectively. We let $h_u(u)$ denote the density of the marginal of H with respect to u, and we define $f_u(u)$ likewise. Further, we let $h_m(m \mid u)$ denote the density of the distribution of m conditional on u according to H. For $(\underline{r}, \overline{r}) \in \mathbb{R}^2$, aggregate welfare is given by

$$\begin{aligned} W(\underline{r}, \overline{r}) &= \gamma \int_{-\infty}^{\infty} l\left(\max\left\{\underline{r}, \min\left\{\overline{r}, u\right\}\right\} - u\right) q(u) du \\ &\quad + (1-\gamma) \int_{-\infty}^{\infty} \left\{ \int_{0}^{\infty} l\left(\max\left\{\underline{r}, \min\left\{\overline{r}, u-m\right\}\right\} - u\right) h_m(m \mid u) dm \right\} h_u(u) du \\ &= \gamma \left[\int_{u \leq \underline{r}} l\left(\underline{r} - u\right) q(u) du + \int_{u \geq \overline{r}} l\left(\overline{r} - u\right) q(u) du \right] \\ &\quad + (1-\gamma) \int_{-\infty}^{\infty} \left\{ \int_{0}^{u-\overline{r}} l\left(\overline{r} - u\right) h_m(m \mid u) dm + \int_{u-\overline{r}}^{u-\underline{r}} l\left(-m\right) h_m(m \mid u) dm \right. \\ &\quad \left. + \int_{u-\underline{r}}^{\infty} l\left(\underline{r} - u\right) h_m(m \mid u) dm \right\} h_u(u) du \end{aligned} \tag{2}$$

Recalling that we interpret $\underline{r} = -\infty$ as the absence of a mandate and $\overline{r} = \infty$ as the absence of a cap, we extend the definition of W to the boundaries of $[-\infty, \infty) \times (-\infty, \infty]$ as follows:

$$\begin{aligned} W(\underline{r}, \infty) &= \gamma \int_{u \leq \underline{r}} l\left(\underline{r} - u\right) q(u) du \\ &\quad + (1-\gamma) \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{u-\underline{r}} l\left(-m\right) h_m(m \mid u) dm + \int_{u-\underline{r}}^{\infty} l\left(\underline{r} - u\right) h_m(m \mid u) dm \right\} h_u(u) du \end{aligned} \tag{3}$$

for $\underline{r} \in \mathbb{R}$,

$$\begin{aligned} W\left(-\infty, \overline{r}\right) &= \gamma \int_{u \geq \overline{r}} l\left(\overline{r} - u\right) q(u) du \\ &\quad + (1-\gamma) \int_{-\infty}^{\infty} \left\{ \int_{0}^{u-\overline{r}} l\left(\overline{r} - u\right) h_m(m \mid u) dm + \int_{u-\overline{r}}^{\infty} l\left(-m\right) h_m(m \mid u) dm \right\} h_u(u) du \end{aligned} \tag{4}$$

for $\overline{r} \in \mathbb{R}$, and

$$W(-\infty,\infty) = (1-\gamma)\int_{-\infty}^{\infty}\left\{\int_{0}^{\infty} l\left(-m\right) h_m(m \mid u)dm\right\} h_u(u)du \tag{5}$$

Given our assumption that $|l|$ has a bounded derivative, we can bound the terms $l\left(\underline{r}-u\right)$ and $l\left(\underline{r}-u\right)$ above and below by linear functions of u, and we can bound $l\left(-m\right)$ above and below by linear functions of m. Because the means of u and m are finite for both G and H (Assumption 2), we can then conclude that $W(\underline{r},\overline{r})$ is finite for all $(\underline{r},\overline{r}) \in [-\infty,\infty) \times (-\infty,\infty]$. Using the preceding expressions and bounds, it is also straightforward to verify that

$$\lim_{\overline{r}\to\infty} W(\underline{r},\overline{r}) = W(\underline{r},\infty)$$

for $\underline{r} \in \mathbb{R} \cup \{-\infty\}$, and

$$\lim_{\underline{r}\to-\infty} W(\underline{r},\overline{r}) = W(-\infty,\overline{r})$$

for $\overline{r} \in \mathbb{R} \cup \{\infty\}$.

Part (i) By applying the Leibniz rule to (2) (or to (4), in case of $\underline{r} = -\infty$), we learn that, for all $(\underline{r},\overline{r}) \in [-\infty,\infty) \times \mathbb{R}$:

$$\begin{aligned}\frac{\partial W(\underline{r},\overline{r})}{\partial \overline{r}} &= \gamma\int_{u\geq\overline{r}} l'\left(\overline{r}-u\right) q(u)du + (1-\gamma)\int_{-\infty}^{\infty}\left\{\int_{0}^{u-\overline{r}} l'\left(\overline{r}-u\right) h_m(m \mid u)dm\right\} h_u(u)du \\ &\geq (1-\gamma)\int_{-\infty}^{\infty} l'\left(\overline{r}-u\right) P_H\left(m < u-\overline{r} \mid u\right) h_u(u)du,\end{aligned}$$

where the strictness of the inequality follows from $l'\left(\overline{r}-u\right) > 0$ for $u > \overline{r}$. If $u < \overline{r}$, then $P_H(m \leq u-\overline{r}|u) \leq P_H(m \leq 0|u) = 0$ by Assumption 1. Moreover, for $u \geq \overline{r}$, we have $l'\left(\overline{r}-u\right) \geq 0$, with strict inequality in the case of $u > \overline{r}$. Therefore the derivative is strictly positive. Recalling in addition that $\lim_{\overline{r}\to\infty} W(\underline{r},\overline{r}) = W(\underline{r},\infty)$ for $\underline{r} \in \mathbb{R} \cup \{-\infty\}$, we see that the Choice Architect does not impose a binding upper bound $\overline{r}$.

Part (ii) We set $\overline{r} = \infty$ (based on part (i)) and, for simplicity, suppress it in the notation. Applying the Leibniz rule to (3), we can write $\frac{dW(\underline{r})}{d\underline{r}}$ as follows:

$$\begin{aligned}\frac{dW(\underline{r})}{d\underline{r}} &= \gamma\int_{u\leq\underline{r}} l'\left(\underline{r}-u\right) q(u)du + (1-\gamma)\int_{-\infty}^{\infty}\left\{\int_{u-\underline{r}}^{\infty} l'(\underline{r}-u) h_m(m \mid u)dm\right\} h_u(u)du \\ &= \int_{u\leq\underline{r}} l'\left(\underline{r}-u\right)\left[\gamma q(u) + (1-\gamma) h_u(u)\right] du \\ &\quad + (1-\gamma)\int_{u>\underline{r}} l'(\underline{r}-u) P_H\left(m \geq u-\underline{r} \mid u\right) h_u(u)du\end{aligned} \tag{6}$$

The first term is strictly negative while the second is strictly positive. It easy to check that, as $\underline{r} \to \infty$, the first term remains bounded away from zero while the second converges to zero. It follows that there is some

finite threshold, r^U, above which W is decreasing in $\underline{r}$.

Endowing the extended real half-line, $[-\infty, r^U]$, with the order topology renders it compact. We have seen that $W(\underline{r})$ is differentiable, and hence continuous (in the order topology), on $\mathbb{R}$. Adding the fact that $\lim_{\underline{r}\to-\infty} W(\underline{r}) = W(-\infty)$, we see that W is continuous on $[-\infty, r^U]$. It therefore follows that the set of maximizers is a non-empty closed subset of $[-\infty, r^U]$, and consequently that the smallest maximizer is well-defined.

Surrogate choices (part (iii)) We begin by writing aggregate welfare as a function of the surrogate choice, $s \in \mathbb{R}$:

$$W(s) = \int_{-\infty}^{+\infty} l\left(s-u\right) f_u(u)du \tag{7}$$

It follows that

$$\begin{aligned}\frac{dW(s)}{ds} &= \int_{-\infty}^{+\infty} l'\left(s-u\right) f_u(u)du \\ &= \int_{u\leq s} l'\left(s-u\right) f_u(u)du + \int_{u\geq s} l'\left(s-u\right) f_u(u)du\end{aligned}$$

The first term is strictly negative, while the second is strictly positive. Because we have assumed that l' is bounded, it follows that the first term converges to zero as $s \to -\infty$ while the second remains bounded away from zero; consequently there exists some finite value s_L such that the derivative is strictly positive for $s < s_L$. Similarly, the second term converges to zero as $s \to \infty$ while the first remains bounded away from zero; consequently there exists some finite value s_U such that the derivative is strictly negative for $s > s_U$. It follows that any optimum on $[s_L, s_U]$ is a global optimum. Because W is continuous in $s \in \mathbb{R}$, the set of maximizers is a non-empty closed subset of $[s_L, s_U]$, and consequently the smallest maximizer is well-defined.

A.2 Proof of Proposition 2

Our claims concerning conventional behavioral welfarism follow directly from the fact that, within our model, the only connection between c^A on the one hand and either mandates or surrogate choices on the other runs through the Choice Architect's beliefs about the distributions of u and m for Choosers. Consequently, the proof focuses on the implications of ideals-projective and mistakes-projective paternalism.

Mandates and caps For the purpose of this proof, we will use $\Phi(x)$ and $\phi(x)$ to denote, respectively, the standard normal CDF and density at x. The marginal distributions for u and z are normal with means and standard deviations of (μ_u, σ_u) and (μ_m, σ_m), respectively. The conditional distributions for u and z are normal with means and standard deviations of $(\hat{\mu}_u(z), \hat{\sigma}_u) = \left(\mu_u + \rho\frac{\sigma_u}{\sigma_m}(z-\mu_m), \sigma_u\sqrt{1-\rho^2}\right)$ and $(\hat{\mu}_m(u), \hat{\sigma}_m) = \left(\mu_m + \rho\frac{\sigma_m}{\sigma_u}(u-\mu_u), \sigma_m\sqrt{1-\rho^2}\right)$. Because we are holding the standard deviations and correlations fixed, we will write the optimal mandate as $\underline{r}^*(\mu_u, \mu_m)$.

Lemma 1

(i) $\underline{r}^*(\mu_u, \mu_m)$ *is increasing in* μ_m.

(ii) $\underline{r}^*(\mu_u, \mu_m)$ *is increasing in* μ_u.

Proof. **Part (i)** We begin by rewriting the objective function (3) to reflect Assumption 4. For clarity, we suppress the dependence of W on $\overline{r}$ (which we know is fixed at ∞), while adding $\mu = (\mu_u, \mu_m)$ as explicit arguments.

$$\begin{aligned} W(\underline{r}, \mu) &= \int_{-\infty}^{\infty} \left\{ \int_{M} l\left(\max\left\{\underline{r}, u - \max\left\{z, 0\right\}\right\} - u\right) \phi\left(\frac{z - \hat{\mu}_m(u)}{\hat{\sigma}_m}\right) \frac{1}{\hat{\sigma}_m} dz \right\} \phi\left(\frac{u - \mu_u}{\sigma_u}\right) \frac{1}{\sigma_u} du \\ &= \int_{\underline{r}}^{\infty} \left\{ \int_{-\infty}^{0} l\left(0\right) \phi\left(\frac{z - \hat{\mu}_m(u)}{\hat{\sigma}_m}\right) \frac{1}{\hat{\sigma}_m} dz + \int_{0}^{u-\underline{r}} l\left(-z\right) \phi\left(\frac{z - \hat{\mu}_m(u)}{\hat{\sigma}_m}\right) \frac{1}{\hat{\sigma}_m} dz \right. \\ &\quad \left. + \int_{u-\underline{r}}^{\infty} l\left(\underline{r} - u\right) \phi\left(\frac{z - \hat{\mu}_m(u)}{\hat{\sigma}_m}\right) \frac{1}{\hat{\sigma}_m} dz \right\} \phi\left(\frac{u - \mu_u}{\sigma_u}\right) \frac{1}{\sigma_u} du \\ &\quad + \int_{-\infty}^{\underline{r}} l\left(\underline{r} - u\right) \phi\left(\frac{u - \mu_u}{\sigma_u}\right) \frac{1}{\sigma_u} du \end{aligned} \tag{8}$$

Applying the Leibniz rule, we obtain:

$$\begin{aligned} \frac{\partial W(\underline{r}, \mu)}{\partial \underline{r}} &= \int_{\underline{r}}^{\infty} l'\left(\underline{r} - u\right) \left[1 - \Phi\left(\frac{u - \underline{r} - \hat{\mu}_m(u)}{\hat{\sigma}_m}\right)\right] \phi\left(\frac{u - \mu_u}{\sigma_u}\right) \frac{1}{\sigma_u} du \\ &\quad + \int_{-\infty}^{\underline{r}} l'\left(\underline{r} - u\right) \phi\left(\frac{u - \mu_u}{\sigma_u}\right) \frac{1}{\sigma_u} du \end{aligned} \tag{9}$$

Now consider $\mu = (\mu_u, \mu_m)$ and $\mu' = (\mu_u, \mu'_m)$ with $\mu'_m > \mu_m$. Then

$$\begin{aligned} \frac{\partial W(\underline{r}, \mu')}{\partial \underline{r}} - \frac{\partial W(\underline{r}, \mu)}{\partial \underline{r}} &= \int_{\underline{r}}^{\infty} l'\left(\underline{r} - u\right) \left[\Phi\left(\frac{u - \underline{r} - \left(\mu_m - \rho \frac{\sigma_m}{\sigma_u}(u - \mu_u)\right)}{\hat{\sigma}_m} \right) \right. \\ &\quad \left. - \Phi\left(\frac{u - \underline{r} - \left(\mu'_m + \rho \frac{\sigma_m}{\sigma_u}(u - \mu_u)\right)}{\hat{\sigma}_m} \right) \right] \phi\left(\frac{u - \mu_u}{\sigma_u}\right) \frac{1}{\sigma_u} du \end{aligned}$$

With $\mu'_m > \mu_m$, the term in square brackets is strictly positive for all u. With $u > \underline{r}$, we have $l'\left(\underline{r} - u\right) > 0$. Thus, $\frac{\partial W(\underline{r}, \mu)}{\partial \underline{r}}$ is strictly increasing in μ_m. It then follows from Topkis' monotone selection theorem that $\underline{r}^*(\mu_u, \mu_m) \leq \underline{r}^*(\mu_u, \mu'_m)$, as claimed.

Part (ii) Rewriting (2) for the special case of $\overline{r} = \infty$ based on our distributional assumptions, we see that for any $k \in \mathbb{R}$, we have

$$W(\underline{r}+k,\mu_u+k,\mu_m)=\gamma\int_{-\infty}^{\infty} l\left(\max\left\{\underline{r}+k,u\right\}-u\right)\phi\left(\frac{u-(\mu_u+k)}{\sigma_u}\right)\frac{1}{\sigma_u}du$$
$$+(1-\gamma)\int_{-\infty}^{\infty}\left\{\int_{0}^{\infty} l\left(\max\left\{\underline{r}+k,u-m\right\}-u\right)\phi\left(\frac{z-\left(\mu_m+\rho\frac{\sigma_m}{\sigma_u}(u-\mu_u-k)\right)}{\hat{\sigma}_m}\right)\frac{1}{\hat{\sigma}_m}dz\right\}$$
$$\times\ \phi\left(\frac{u-(\mu_u+k)}{\sigma_u}\right)\frac{1}{\sigma_u}du$$

Now we define $v=u-k$. With a change of variables from u to v, we can then write the welfare function as:

$$W(\underline{r}+k,\mu_u+k,\mu_m)=\gamma\int_{-\infty}^{\infty} l\left(\max\left\{\underline{r},v\right\}-v\right)\phi\left(\frac{v-\mu_u}{\sigma_u}\right)\frac{1}{\sigma_u}dv$$
$$+(1-\gamma)\int_{-\infty}^{\infty}\left\{\int_{0}^{\infty} l\left(\max\left\{\underline{r},v-m\right\}-v\right)\phi\left(\frac{z-\left(\mu_m+\rho\frac{\sigma_m}{\sigma_u}(v-\mu_u)\right)}{\hat{\sigma}_m}\right)\frac{1}{\hat{\sigma}_m}dz\right\}\phi\left(\frac{v-\mu_u}{\sigma_u}\right)\frac{1}{\sigma_u}dv$$
$$=W(\underline{r},\mu_u,\mu_m)$$

It follows that increasing μ_u by k simply shifts the entire welfare function rightward by k, which means any maximizer increases by k. The same statement obviously applies to the smallest maximizer. □

Lemma 1 immediately yields the following corollary:

Corollary 1 *For any (μ_u,μ_m) and (μ'_u,μ'_m) with $\mu'_u\geq\mu_u$ and $\mu'_m\geq\mu_m$ and at least one inequality strict, we have $\underline{r}^*(\mu'_u,\mu'_m)>\underline{r}^*(\mu_u,\mu_m)$.*

Proof. We can decompose a change from (μ_u,μ_m) to (μ'_u,μ'_m) into two steps: (a) (μ_u,μ_m) to (μ_u,μ'_m); (b) (μ_u,μ'_m) and (μ'_u,μ'_m). Applying Lemma 1 to each of these steps delivers the desired conclusion.

□

The proposition concerns the distribution of $\underline{r}^*$ conditional on c^A, the selection Choice Architects make for themselves. We use the preceding property to show that this distribution increases (in the sense of FOSD) as we increase the Choice Architects' consumption level from c^A to any $c^{A'}>c^A$.

Let M_u denote the random variable with distribution $G_{u^A}(\cdot|c^A)$ and M'_u denote the random variable with distribution $G_{u^A}(\cdot|c^{A'})$. Define M_m and M'_m likewise. Assumption 3 implies that $G_{u^A}(\cdot|c^{A'})$ first-order dominates $G_{u^A}(\cdot|c^A)$, and $G_{m^A}(\cdot|c^A)$ first-order dominates $G_{m^A}(\cdot|c^{A'})$. By Strassen's monotone coupling theorem, there exists an underlying state space Ω such that we can write all of these random variables as functions of the state ω, with $M'_u(\omega)\geq M_u(\omega)$ and $M'_m(\omega)\leq M_m(\omega)$, where in each case the inequality is strict for a set of states with positive measure.

By the definition of ideals-projection, there exists a function $\varphi:\mathbb{R}\to\mathbb{R}^2$ such that $(\mu_u,\mu_m)=\varphi(u^A)$ with φ strictly increasing in the first argument and weakly increasing in the second. For any given state

ω we have $M'_u(\omega) \geq M_u(\omega)$, which according to the Corollary implies $\underline{r}^*\left(\varphi\left(M'_u\left(\omega\right)\right)\right) \geq \underline{r}^*\left(\varphi\left(M_u\left(\omega\right)\right)\right)$, with strict inequality for a set of states with positive measure. Thus, with ideals-projective paternalism, the distribution of $\underline{r}^*$ conditional on $c^{A'}$ first-order dominates the distribution of $\underline{r}^*$ conditional on c^A.

By the definition of mistakes-projection, there exists a function $\psi : \mathbb{R} \to \mathbb{R}^2$ such that $(\mu_u, \mu_m) = \psi(m^A)$ with ψ weakly increasing in the first argument and strictly increasing in the second. For any given state ω we have $M_u(\omega) \geq M'_u(\omega)$, which according to the Corollary implies $\underline{r}^*\left(\varphi\left(M_u\left(\omega\right)\right)\right) \geq \underline{r}^*\left(\varphi\left(M'_u\left(\omega\right)\right)\right)$, with strict inequality for a set of states with positive measure. Thus, with mistakes-projective paternalism, the distribution of $\underline{r}^*$ conditional on c^A first-order dominates the distribution of $\underline{r}^*$ conditional on $c^{A'}$.

Surrogate choices Because we are holding the standard deviations and correlations fixed, we will write the optimal surrogate choice as $s^*(\mu_u, \mu_m)$.

Lemma 2

(i) $s^*(\mu_u, \mu_m)$ *is independent of* μ_m.

(ii) $s^*(\mu_u, \mu_m)$ *is increasing in* μ_u.

Proof. Part (i) follows immediately from equation (7), which implies that aggregate welfare is independent of μ_m.

For part (ii), we begin by writing aggregate welfare as a function of the surrogate choice, $s \in \mathbb{R}$, under our distributional assumptions:

$$W(s, \mu) = \int_{-\infty}^{+\infty} l\left(s - u\right) \phi\left(\frac{u - \mu_u}{\sigma_u}\right) \frac{1}{\sigma_u} du \tag{10}$$

Note that for any $k \in \mathbb{R}$, we have

$$W(s + k, \mu_u + k, \mu_m) = \int_{-\infty}^{+\infty} l\left(s + k - u\right) \phi\left(\frac{u - (\mu_u + k)}{\sigma_u}\right) \frac{1}{\sigma_u} du$$

Now we define $v = u - k$. With a change of variables from u to v, we can then write the welfare function as:

$$W(s + k, \mu_u + k, \mu_m) = \int_{-\infty}^{+\infty} l\left(s - v\right) \phi\left(\frac{v - \mu_u}{\sigma_u}\right) \frac{1}{\sigma_u} dv = W(s, \mu_u, \mu_m)$$

It follows that increasing μ_u by k simply shifts the entire welfare function rightward by k, which means any maximizer increases by k. The same statement obviously applies to the smallest maximizer.

□

The rest of the proof is essentially the same as that of Proposition 2, with $s^*(\mu_u, \mu_m)$ replacing $\underline{r}^*(\mu_u, \mu_m)$. The one material difference is that $s^*(\mu_u, \mu_m)$ is independent of μ_m. As a result, with mistakes-projective paternalism, we can only say that the distribution of s^* conditional on c^A *weakly* first-order dominates the distribution of $\underline{r}^*$ conditional on $c^{A'} > c^A$.

B Additional Analysis

B.1 The Induced Chooser Preferences and Exogenous Removal conditions

Here, we detail the Induced Chooser Preferences and Exogenous Removal conditions summarized in Section 3.5.

The Induced Chooser Preferences condition The induced preference parameter $\hat{\delta}_i$ equals 0, 0.5, or 1, each with equal probability. These parameter values are chosen to minimize complexity. Choice Architects are aware of this distribution, but no Choice Architect knows the specific value of $\hat{\delta}_i$ that the Chooser assigned to her will face. We describe options as bundles (X, Y) of 'gold tokens' and 'silver tokens,' respectively, to be exchanged for Euros the day of the experiment. Choice Architects know that the Chooser will be able to exchange each gold token for one Euro, while he will receive $\hat{\delta}_i$ for each silver token.

In the Induced Chooser Preferences condition, the frequency with which Choice Architects exclude at least one option from a menu equals 22.6% (s.e. 2.1 percentage points). The corresponding number for decisions in the Main condition, by comparison, is much higher, 39.8% (s.e. 2.3 percentage points). As mentioned in the main text, the monotonic relationship between larger values of X and removal frequencies that characterizes the Main condition vanishes entirely. Removal rates for the options in the Induced Chooser Preferences condition with the largest, middle, and smallest values of X are 12.9% (s.e. 1.6%), 5.1% (s.e. 1.1%), and 7.8% (s.e. 1.3%), respectively.[1] Moreover, across all rounds and Choice Architects in the Main condition, subjects indicate in 29.6% (s.e. 2.2%) of cases that the unrestricted opportunity set is worse for the Chooser. This number drops to 15.3% (s.e. 1.8%) in the Induced Chooser Preferences condition. Similarly, the elicited compensating variations are more than twice as large in the Main condition as in the Induced Chooser Preferences condition: 0.079 (s.e. 0.015) versus 0.038 (s.e. 0.011).

The Exogenous Restriction condition As explained in the main text, in one round of this condition, Choice Architects decide between making all options available, removing the least patient option, and removing the two least patient options, whereas in another round of this condition, we exogenously remove the least patient option. We find that the availability frequency for the middle option falls slightly from 89.7% when the least restrictive opportunity set contains three options, to 85.8% when we remove the least patient option exogenously. One would expect to see some decline merely as a consequence of noisy choice. For example, if Choice Architects randomize between all their options with equal probabilities, then they would make the middle option available two-thirds of the time when all three options are available, versus one-half of the time when only two options are available. As explained in the main text, we control for the effects of noisy behavior by including two additional rounds that are identical to the first pair, except that the default is the most restrictive opportunity set, consisting of just the most patient option. By taking action, the Choice Architect adds rather than removes options. Random choosers will add options, but subjects who are motivated by the desire to restrict others' opportunity sets will not. In these two rounds, we find that the exogenous removal of the least patient option causes a similar decline in the availability of the middle

[1] The higher removal rate for the first option may reflect a variety of motives, including fairness concerns. A Chooser facing $\hat{\delta}_i = 0$ can obtain no more than €4 in the Induced Chooser Preferences condition. A Chooser facing $\hat{\delta}_i = 1$, in contrast, obtains €15 from choosing the first option. By removing that option, the Choice Architect can reduce inequality across Choosers.

option (76.2% and 73.9%). The difference-in-differences is far from statistically significant in a regression with session and order fixed effects ($p > 0.6$).

The foregoing analysis is based on within-subject variation. The results do not change qualitatively if we restrict attention to cross-subject variation by only including each Choice Architect's first decision from among these four rounds. The resulting estimates of the decline in the frequency with which Choice Architects offer the middle option are virtually the same regardless of whether we specify the most restrictive set or the least restrictive set as the default option: it falls from 89.0% to 81.7% in the first instance, and from 78.3% to 71.1% in the second ($p > 0.9$ for the difference-in-differences).

We conclude that there is no evidence indicating that Choice Architects are primarily motivated by a desire to exert control, even when they can potentially rationalize such action as beneficial.

B.2 Assessing anchoring and demand for consistency

Mechanisms such as anchoring or a demand for consistency potentially create an artificial relation between the Choice Architect's own choices in stage 2 of the laboratory component and the opportunity sets she has previously constructed for Choosers. We test for such mechanisms using the following strategy. We consider the variable d_i, defined as the difference between a Choice Architect's percentile rank of her own choices in stage 2 and the inverse of her patience rank in the online component.[2] If the confounding mechanisms are quantitatively important, then we should observe a positive relation between mandates and d_i: if mandates influence a Choice Architect's own choices in stage 2 (through anchoring or a demand for consistency), then subjects who impose stricter mandates should exhibit larger discrepancies between their patience rank in the online component and Stage 2 of the laboratory component. (This inference rests on the assumption that the confounding mechanisms do not apply to the relation between mandates and choices in the online component.)

Formally, we regress d_i on the mean mandate a Choice Architect imposes in the Main condition (without front-end delay). We obtain a coefficient estimate of 0.03 (s.e. 0.06). This relation is far from statistically significant ($p = 0.62$). We conclude that there are no quantitatively important confounding mechanisms that artificially relate Choice Architects' mandates to the choices they make for themselves in stage 2.

B.3 Attention test

Table A1 displays the attention test administered in Stage 2 of the online component. For each statement, the subject indicates whether it is true or false. The table also lists the percentage of correct responses next to each statement. The mean test score is 6.40 out of 8, with a standard deviation of 1.23. The lowest performance is on question 5, which was answered correctly by 58% of participants. This question concerns the Induced Chooser Preference condition. Question 2, on front-end delay, was also answered incorrectly relatively often, in spite of the fact that the information on front-end delay was presented prominently (though only in a single round of Stage 1). While some subjects might have forgotten this detail when

[2] Specifically, we first calculate the mean amount of money they choose to receive immediately in stage 2, averaging across the three-option menus used in the Main condition (without front-end delay). We obtain each Choice Architect's percentile rank according to that variable. We define the variable d_i as the difference between the percentile rank just defined and the inverse percentile rank of a Choice Architect's patience elicited in the online component.

taking this test, the fact that the introduction of a front-end delay leads to significant behavioral differences suggests they did pay attention. Indeed, if we only include subjects who correctly answered the test question on front-end delay, the coefficients of the regressions corresponding to columns 1 to 3 of Table 4 are given by -0.28 (s.e. 0.12), -0.20 (s.e. 0.07), and -0.11 (s.e. 0.13), respectively. If we only include subjects who did not answer that question correctly, the coefficients barely change; they are given by -0.35 (s.e. 0.15), -0.28 (s.e. 0.09), and -0.12 (s.e. 0.15).

Table A1: Attention test concerning decisions made for Choosers.

1. In some rounds I could remove options, but I could never add options for the future experiment participant. (False, 85%)
2. For all options which the future experiment participant could receive, the early payment will always be on the day of the experiment. (False, 60%)
3. For some options that the future experiment participant could have received, the late payment will occur up to 7 months after the day of the study. (False, 86%)
4. In some rounds, I had to make a single option available to the future experiment participant; I could not make multiple options available, even if I wanted to. (True, 77%)
5. Some rounds concerned gold tokens and silver tokens. These rounds only concerned money that would be paid soon after the experiment, not money that would be paid only months after the experiment. (True, 58%)
6. Other experiment participants could receive €0.40, €0.50, or €0.60 per silver token, each with the same probability. (False, 98%)
7. In some lines of the decision lists, the base payment of the future participant could be increased. In other rounds it could be decreased. (True, 93%)
8. If I made some options unavailable, this means that the other experiment participant cannot see that option, and therefore does not need to think about these options. (False, 83%)

Note: Fractions of correct answers are given in parentheses.

B.4 Projective Paternalism and Information treatments

Here, we address the question whether informational interventions can affect the extent of projective paternalism. Ex ante, the answer is not obvious. On the one hand, Section 3.4 shows that interventions respond substantially to Chooser-specific information. On the other hand, previous research on the false consensus effect documents its resistance to informational interventions (Krueger and Clement, 1994; Engelmann and Strobel, 2012).

We examine two different informational interventions. First, we test whether the information provided in the Chooser Information condition affects the extent of projective paternalism. To this end, we regress mandates on Choice Architects' patience percentiles using data from the Chooser Information Treatment. We allow the coefficient of Choice Architects' patience percentile to vary freely across the four Chooser statements. As in the main text, we focus on non-libertarian Choice Architects. We control for session, order, and menu fixed effects, and we cluster standard errors by subject. Column 1 of Table A2 displays the results. The estimated coefficients are all similar, and we cannot reject the hypothesis that they are identical, either in pairwise comparisons or in a joint test ($p > 0.8$ for the joint test, $p > 0.4$ for each pairwise test). In column 2, we use beliefs as dependent variable. This regression allows us to examine the effect of the Chooser statements on the false consensus effect. We find that neither of the Chooser statements

significantly affect the magnitude of the false consensus effect ($p > 0.5$ for the joint test, $p > 0.15$ for each pairwise test).

There are two explanations for the foregoing finding. First, it is possible that projective paternalism is a robust phenomenon with respect to the provision of information. Second, the verbal statements about Chooser's inclinations and subjective experience require interpretation, which might be influenced by Choice Architects' own preferences in a way that leads to projective paternalism.

Table A2: Projective paternalism and information provision

VARIABLES	(1) Mandate	(2) Belief	(3) Mandate	(4) Belief
Patience %-ile				
× *Chooser*				
patient, happy	-0.915**	-1.313***		
	(0.358)	(0.165)		
impatient, happy	-0.798**	-0.947***		
	(0.353)	(0.235)		
patient, unhappy	-0.675**	-1.132***		
	(0.332)	(0.193)		
impatient, unhappy	-1.022***	-1.027***		
	(0.387)	(0.210)		
× *Choice dist. info.*			-0.290	-1.022***
			(0.374)	(0.253)
× *Not in Choice Dist. info.*			-1.048***	-1.280***
			(0.269)	(0.170)
Choice dist. info			-0.334	-0.183
			(0.327)	(0.200)
Mean of dep. var.	2.825	1.505	3.085	1.442
	(0.082)	(0.059)	(0.075)	(0.051)
p of joint test for differences in slopes	0.860	0.581	0.107	0.393
Observations	716	716	732	729
Number of subjects	179	179	244	243

Notes: Method: OLS. Unit of observation: subject-round pairs. Dependent variables: *Mandate* is the maximum amount the Choice Architect allows the Chooser receive immediately (as in Section 3.3). *Belief* is the Choice Architect's beliefs about the mean amount a Chooser will receive early if allowed to choose without restrictions. Controls: *Patience %-ile* is the Choice Architect's percentile rank according to the average number of months she is willing to delay the receipt of the larger payment in the online tasks. Other controls include session, order, and menu fixed effects. Samples: All regressions limited to non-libertarian subjects who responded monotonically to all multiple-decision lists in the online component. Columns 1 and 2 are based on the Chooser Information condition using subjects who are not in the Choice Distribution Information treatment. Columns 3 and 4 are based on the Main condition (without front-end delay) adding subjects from the Choice Distribution Information treatment. Standard errors: clustered at the subject level. $^{*}p < 0.1,^{**}p < 0.05,^{***}p < 0.01$.

To evaluate the second interpretation, we study data from 100 additional Choice Architects who participated in the *Choice Distribution Information* condition. The experiment for these Choice Architects proceeded in the same fashion as for all other Choice Architects, with the exception that, in each round, these Choice Architects could click a button to view the distribution of previous selections Choosers made when all options were available. Upon clicking the button they observed, for each of the three options in the menu, a line of text of the following form: *Out of 100 previous Choosers, N choose X today and Y in six months.*[3]

[3]We calibrated the numbers N based on online pilot studies. Each Choice Architect in each session session faced a 25% chance of

In this analysis, we focus again on non-libertarian Choice Architects. The fraction of libertarians in the Choice Distribution Information treatment is 32%, compared to 38.3% libertarians amongst subjects not in that treatment ($p > 0.25$).

Choice Architects' decisions to reveal the information are instructive: 39.7% of non-libertarian subjects in the Choice Distribution information treatment never viewed the information provided, even though they could do so costlessly, at the click of a button. The remaining subjects viewed the information in 46% of all rounds, on average.

To study the effect of distributional information formally, we regress mandates on Choice Architects' patience percentile using data from the Main condition (without front-end delay). We allow the coefficient to vary freely depending on whether a Choice Architect is in the Choice Distribution Information treatment, and we include an indicator for that treatment. We control for session, order, and menu fixed effects, and we cluster standard errors by subject.

Column 3 of Table A2 displays the results. Projective paternalism appears to be attenuated in the Choice Distribution Information treatment. However, given the high frequency with which subjects choose not to view the information, we interpret this result with caution. Two further results suggest a cautious interpretation. First, the difference in estimated coefficients across the two treatments is not statistically significant ($p > 0.1$). Second, column 4 shows the estimated coefficients when we replace the dependent variable with Choice Architects' beliefs about unrestricted Choosers' choices. We find that the effect of the Choice Distribution Information treatment is much weaker and statistically insignificant ($p > 0.39$). Because the Choice Distribution information treatment does not alter beliefs, we consider it unlikely that the treatment has a genuine effect on mandates.

Overall, we conclude that ideals-projective paternalism is robust to the kind of information we have provided.

B.5 Choosers' choices

Table A3 lists the choices by Choosers. Each Chooser ranked the three options associated with single round of the experiment. The modal Chooser ranks the most patient option first and the least patient option last, though each option is ranked first by some Chooser. Choosers participate merely to incentivize the Choice Architects; we do not draw inferences from their behavior. We recorded the minimum data required to determine each Chooser's payment.

being assigned to the Choice Distribution Information condition.

Table A3: Choosers' choices.

	Preference rank		
	First	Second	Third
Money over time ($n = 103$)			
Most patient choice	79.61%	11.65%	8.74%
Middle Choice	14.56%	82.52%	2.91%
Least Patient Choice	5.83%	5.83%	88.35%
Induced preferences ($n = 21$)			
Most patient choice	57.14%	23.81%	19.05%
Middle Choice	23.81%	71.43%	4.76%
Least Patient Choice	19.05%	4.76%	76.19%

Notes: The table shows, for each option, the fraction of Choosers who placed that option at the indicated preference rank position. *Money over time* concerns choices corresponding to the Main, Exogenous Restriction, and Chooser Information conditions. *Induced preference* concerns choices corresponding to the Induced Chooser Preferences condition.

B.6 Price sensitivity

Here, we document the sensitivity of subjects' interventions to the monetary cost of impatience. We begin by regressing an a binary variable indicating whether an option is withheld on the amount of Euros the Choice Architect forfeits by choosing that option rather than the most patient option (which always pays a total amount of €15), and on indicators for whether the option is the least patient or middle option in the menu. For this purpose, the unit of observation is a Choice Architects' decision concerning a given option in a given round. We include observations from the Main treatment without front-end delay, we control for session and order fixed effects, and we cluster standard errors on the subject level. Column 1 of Table A4 displays the result. The coefficient estimates indicate that the magnitude of the monetary cost rather than the relative rank of an option in the choice set is the primary driver of decisions to withhold options. As reported in the main text, a €1 reduction in total payoff increases the probability that an option is withheld by 2.9 percentage points ($p < 0.01$). To reinforce this point, Columns 2 performs the regression including only the middle option of each choice set, which leads to a coefficient of 2.1 percentage points. Column 3 includes only the least patient option of each choice set and yields a coefficient of 10.5 percentage points.

Table A4: Price sensitivity.

VARIABLES	(1) Option withheld	(2) Option withheld	(3) Option withheld
Rank of included option	All	Middle	Least patient
Cost of impatience	0.029***	0.021***	0.105***
	(0.006)	(0.007)	(0.037)
Option rank			
Middle	-0.008		
	(0.021)		
Least patient	0.019		
	(0.059)		
Observations	2,727	909	909
#subj	303	303	303

Notes: Method: OLS. Unit of observation: subject-round-option triplet. Controls: *Cost of impatience* is the amount of Euros a Chooser forfeits by selecting a given option, compared to the most patient option. *Option rank* refers to the placement of an option within a menu. Other controls include session and order fixed effects. Sample: all rounds from the Main condition without front-end delay.

B.7 Beliefs about Chooser welfare

Table A5 displays the effects of - delay and of Chooser-specific information on judgments about the welfare effects of withholding options, as referenced in Sections 3.3 and 3.4, respectively. Table A6 presents regression results concerning beliefs about the welfare effects of removing options using ordered probit specifications. All coefficient estimates of interest retain their signs as well as their statistical significance.

Table A5: Subjects' beliefs about the welfare effects of withholding options in the case of front-end delay and in the Chooser Information condition

Dependent variable	(1) Belief smaller opportunity set better for Chooser	(2)	(3) Negative of Compensating Variation	(4)
Smaller set	Chosen	Chosen	Chosen	Chosen
Menus	1 - 2	5	1 - 2	5
Menu with FED included	Yes	-	Yes	-
Summary statistics for dependent variables				
Distribution of beliefs				
Evaluation opportunity set is				
better	0.297	0.356		
same	0.611	0.540		
worse	0.092	0.103		
Mean negative CV			0.021	-0.224
Distribution of beliefs if options withheld				
Evaluation opportunity set is				
better	0.674	0.693		
same	0.090	0.099		
worse	0.236	0.207		
Mean negative CV			0.095	0.078
Regression results				
Front-end delay	0.077**		0.045*	
	(0.037)		(0.025)	
Chooser information				
Patient, happy		0.077*		0.002
		(0.041)		(0.028)
Patient, unhappy		0.015		-0.029
		(0.040)		(0.025)
Impatient, unhappy		0.116***		0.010
		(0.040)		(0.025)
Observations	606	1,212	583	1,156
Number of subjects	303	303	294	298

Notes: The top half of the table provides summary statistics for the dependent variables, and the bottom half presents regression results. Unit of observation: subject-round pair. Method: Columns 1 and 2, OLS; columns 3 and 4, interval regression. Dependent variables: The dependent variable for columns 1 and 2 measures whether the Choice Architect considers the smaller opportunity set better, equally good, or worse for the Chooser than the unrestricted set, coded as 1, 0, and -1, respectively. For columns 3 and 4, it is the negative of the Choice Architect's beliefs about the compensating variation (CV) of reducing the opportunity set. The smaller opportunity set is the one the Choice Architect has constructed. Controls: All regressions control for altruism and spite, and include session, order, and menu fixed effects. Samples: Each of columns 1 and 2, and each of columns 3 and 4, employs data from a different condition, as indicated. In the bottom panel, each column corresponds to a separate regression. Columns 3 and 4 exclude subjects with multiple switches in the CV elicitation. Standard errors: clustered by subject. $^{*}p < 0.1$,$^{**}p < 0.05$,$^{***}p < 0.01$.

Table A6: Ordered probit regressions corresponding to Tables 3 and A5

	(1)	(2)	(3)	(4)
Smaller set	Chosen	Exogenous	Chosen	Chosen
# options withheld				
1	0.634***			
	(0.144)			
2	1.035***			
	(0.247)			
Mean # options withheld		1.081***		
in Main condition		(0.139)		
Opportunity set contains		-0.308**		
most patient option only		(0.138)		
Front-end delay			0.155**	
			(0.072)	
Chooser information				
Patient, happy				0.142*
				(0.076)
Patient, unhappy				0.028
				(0.073)
Impatient, unhappy				0.217***
				(0.074)
Category thresholds oProbit				
Worse vs. same	-0.535*	0.453	-1.113***	-0.986***
	(0.307)	(0.371)	(0.357)	(0.283)
Same vs. better	1.385***	0.801**	0.815**	0.691**
	(0.310)	(0.369)	(0.351)	(0.279)
Observations	909	606	606	1,212
Number of subjects	303	303	303	303

Notes: The dependent measures whether the Choice Architect considers the smaller opportunity set better, equally good, or worse for the Chooser than the unrestricted set. We assign higher index numbers to more positive beliefs. Unit of observation: subject-round pair. Method: Ordered probit. Otherwise, specifications are the same as in the corresponding OLS regressions in Tables 3 and A5 $^{*}p < 0.1$,$^{**}p < 0.05$,$^{***}p < 0.01$.

B.8 Projective paternalism, beliefs, and measurement error

Here, we examine the possible effects of measurement error in our analysis on the role of beliefs for projective paternalism—specifically Column 2 in Table 7. We do so by instrumenting for the Choice Architect's patience (measured in the online component) using the choices she makes for herself in stage 2 of the laboratory component using all rounds corresponding to the Main condition. We also instrument for the Choice Architect's beliefs about choices corresponding to the Main condition by the same subjects' beliefs about choices corresponding to the Chooser Information and Exogenous Restriction conditions. Aside from the 2SLS estimation method, our specification is the same as in column 2 of Table 7.

Table A7 shows the results. For comparison, column 1 replicates the OLS regression featured in the main text. The specification in Column 2 instruments both predictor variables. We find that coefficient on beliefs more than doubles, while the coefficient on the Choice Architects' patience flips sign. The latter coefficient is measured with such a large standard error that its confidence interval ranges from -1.28 to 2.26. This confidence interval includes both the coefficient estimate from a regression without beliefs (column 1 of Table 7), and the coefficient estimate of the OLS regression in column 1. For completeness, Column 3 presents

a version of this regression in which we only instrument for beliefs, but not for patience, while column 4 only instruments for patience but not for beliefs. We find a substantial but noisily estimated effect of own patience in the latter specification.

Overall, we conclude that beliefs about others' choices play an important role in determining mandates. The estimates of the effect of patience conditional on beliefs, in contrast, are too noisy for us to reach definitive conclusions.

Table A7: Role of beliefs in projective paternalism: instrumental variables regressions

VARIABLES	(1) Mandate	(2) Mandate	(3) Mandate	(4) Mandate
Method	OLS	IV	IV	IV
Patience %-ile	-0.540*	0.491	0.008	-1.149
	(0.318)	(0.903)	(0.384)	(0.832)
Belief %-ile	-1.136***	-2.657***	-2.374***	-0.781
	(0.365)	(0.758)	(0.506)	(0.616)
Instrumented				
Patience %-ile	No	Yes	No	Yes
Belief %-ile	No	Yes	Yes	No
Observations	537	537	537	537
Number of subjects	179	179	179	179

Notes: Unit of observation: subject-round pairs. Dependent variable: *Mandate* is the maximum amount the Choice Architect allows the Chooser receive immediately (as in Section 3.3). Controls: *Patience %-ile* is the Choice Architect's percentile rank according to the average number of months she is willing to delay the receipt of the larger payment in the online tasks. *Beliefs %-ile* is the Choice Architects' percentile rank in terms of her beliefs about the Chooser's patience, as measured by the inverse of the mean amount the Chooser will receive immediately if allowed to choose without restrictions. Additional controls include session, order, and menu fixed effects. Samples: Regressions limited to non-libertarian subjects who responded monotonically to all multiple-decision lists in the online component. Rounds from the Main condition without front-end delay. Standard errors: clustered at the subject level. $^{*}p < 0.1,^{**}p < 0.05,^{***}p < 0.01$.

B.9 Choice Architects' own choices

In Stage 2, Choice Architects choose one of the options from each menu in Panel A of Table 1. Table A8 lists the distribution of these choices for menus corresponding to the Main condition.

Table A8: Choices that Choice Architects make for themselves

Menus	Percent chosen			Obs.	Subjects
	Most patient	Middle	Least patient		
1, 2, 3, 4	78.3	13.0	8.4	1612	403
	(1.7)	(1.2)	(1.2)		
1	68.7	18.5	12.3	195	195
	(3.3)	(2.8)	(2.4)		
2	77.9	9.6	12.5	208	208
	(2.9)	(2.0)	(2.3)		
3	70.7	23.1	6.0	403	403
	(2.3)	(2.1)	(1.2)		
4	87.6	4.7	7.4	403	403
	(1.6)	(1.1)	(1.3)		
1, 2	81.4	10.4	7.9	403	403
Front-end delay	(1.9)	(1.5)	(1.3)		
1, 2	73.4	13.9	12.4	403	403
No front-end delay	(2.2)	(1.7)	(1.6)		

Notes: Standard errors in parentheses.

B.10 Judgments about welfare effects of actual paternalistic policies

Here, we provide material supplementary to our analysis in Section 5 concerning judgments about actual paternalistic policies. Panel A of Table A9 shows the distribution of welfare judgments by policy. Panel B.1 displays the coefficients of ordered probit versions of the regressions concerning the relation between welfare judgments about actual policies and welfare judgments in the experimental decisions (see Panel B.2 in Table 8). As in the main text, we find highly statistically significant relations for all tax policies, but not for restrictions on short-term, high-interest lending. Panel B.2 displays corresponding results when we use beliefs about the negative compensating variation for removing options in the experimental tasks. Here, we find statistically significant coefficient estimates for both tobacco taxes and for restrictions on short-term, high-interest lending, but not for the remaining policies. For all policies, we find the expected positive sign. The lack of statistical significance for two of the tax policies is possibly due to the complex elicitation procedure through which subjects reveal beliefs about compensating variations.

B.11 Robustness check concerning judgments of actual paternalistic policies

The regressions in Table 8 include an array of control variables. Here, we present estimates of the same regressions in which we only retain session and order fixed effects as control variables. As Panel A.1 of Table A10 shows, in the case without controls, the relation between mandates in the experimental decisions and judgments of real-world paternalistic policies is slightly weaker for each policy. Even though the sign of the estimated coefficients remains unchanged and the magnitude remains similar in each case, this attenuation causes a loss in statistical significance for three of the four policies. Importantly, the relation to the average of

Table A9: Experimental decisions and support for real-world paternalistic policies.

Policy	(1) Increase alcohol tax	(2) Increase tobacco tax	(3) Introduce sugary drinks tax	(4) Tighten restrictions on short-term lending
	A. Summary statistics for dependent variables			
Dist. of beliefs about welfare effects				
significantly worse off	0.112	0.050	0.050	0.072
a little worse off	0.251	0.099	0.166	0.243
neither better nor worse off	0.228	0.176	0.310	0.357
a little better off	0.270	0.295	0.330	0.256
significantly better off	0.139	0.380	0.144	0.072
	B. Relation between beliefs about welfare effects in the laboratory and concerning actual policies			
	B.1 Predictor variable: non-incentivized statement			
Average belief exog. smaller choice set better for Chooser	0.205***	0.230***	0.217***	0.077
	(0.078)	(0.083)	(0.076)	(0.076)
Category thresholds oProbit				
1	-1.974***	-2.511***	-3.199***	-0.954
	(0.685)	(0.697)	(0.714)	(0.689)
2	-1.031	-1.851***	-2.271***	0.099
	(0.677)	(0.687)	(0.711)	(0.685)
3	-0.390	-1.214*	-1.361*	1.063
	(0.674)	(0.679)	(0.708)	(0.687)
4	0.537	-0.390	-0.303	2.103***
	(0.675)	(0.678)	(0.708)	(0.695)
Observations	403	403	403	403
	B.2 Predictor variable: belief about compensating variation			
Mean negative CV	0.201	0.379**	0.189	0.214*
	(0.138)	(0.152)	(0.141)	(0.130)
Category thresholds oProbit				
1	-1.791**	-2.491***	-3.178***	-0.889
	(0.710)	(0.710)	(0.741)	(0.716)
2	-0.861	-1.824***	-2.276***	0.189
	(0.703)	(0.701)	(0.736)	(0.713)
3	-0.214	-1.191*	-1.383*	1.129
	(0.701)	(0.692)	(0.732)	(0.716)
4	0.698	-0.375	-0.319	2.181***
	(0.702)	(0.691)	(0.733)	(0.724)
Observations	391	391	391	391

Notes. Method: ordered probit. Unit of observation: subject. Dependent variables: *Beliefs about welfare effects* is the belief about the welfare effect of various policies. Controls: *Average belief exog. smaller choice set better for Chooser* (as in Section 3.2) encodes whether the Choice Architect considers the smaller opportunity set better, equally good, or worse for the Chooser than the unrestricted set, coded as 1, 0, and -1, respectively, averaged across the four rounds of the Exogenous Restriction Condition. Controls: *Average belief exog. smaller choice set better for Chooser* measures whether the Choice Architect considers the smaller opportunity set better, equally good, or worse for the Chooser than the unrestricted set, coded as 1, 0, and -1, respectively, averaged across rounds in the Exogenous Restriction condition. *Mean negative CV* is the negative of the Choice Architect's beliefs about the compensating variation (CV) of reducing the opportunity set, averaged across rounds in the Exogenous Restriction condition. All regressions control for gender, age, self-reported political attitudes, log monthly expenses, high school GPA, university faculty at which the subject's major field of study is offered, weekly alcohol consumption, log days of binge drinking per year (defined as the consumption of at least 4 units of alcohol for females, 5 for males, within a period of two hours (National Institutes on Alcohol Abuse and Alcoholism, 2018)), smoking status, number of cigarettes smoked per day, body mass index, credit card debt, for having taken a short term loan, as well as for session and order fixed effects. For variables measured with interval precision, we use midpoints for analysis. For control variables that subjects chose not to disclose, the regressions impute population means, and include indicators for whether a variable's value was missing. Samples: includes 303 subjects who participated in the main experiment, plus 100 subjects who participated in the Choice Distribution Information Condition (see Appendix B.4). Panel B.2 excludes subjects with multiple switches in the elicitation of beliefs about compensating variations. Standard errors: clustered by subject. $^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$.

judgments across the four policies remains statistically significant. In Panel A.2 we find a slight attenuation of the estimated coefficients concerning welfare beliefs. Here, the statistical significance of the coefficient estimates remains unchanged for each policy. Panel B shows the effect of excluding subject-specific control variables on the estimates of projective paternalism concerning real-world policies. Here, we find that the magnitude and levels of statistical significance remain highly similar to the case that includes the controls.

Table A10: Replication of estimates in Table 8 without subject-specific control variables.

Policy	(1) All	(2) Increase alcohol tax	(3) Increase tobacco tax	(4) Introduce sugary drinks tax	(5) Tighten restrictions on short-term lending
	A. Relation between laboratory choice and policy attitudes				
	A.1 Dependent variable: support for policy proposal				
Average mandate imposed on Chooser (max. € paid immediately)	-0.091**	-0.102	-0.101**	-0.083	-0.078
	(0.040)	(0.069)	(0.051)	(0.066)	(0.052)
Observations	403	403	403	403	403
	A.2 Dependent variable: beliefs about welfare effect of policy proposal				
Average belief exog. smaller choice set better for Chooser	0.176***	0.200**	0.222***	0.198***	0.085
	(0.052)	(0.084)	(0.080)	(0.072)	(0.075)
Observations	403	403	403	403	403
	B. Projective paternalism with actual policies				
Alcohol consumption					
Alcohol units / week		-0.059*			
		(0.033)			
log(days binge drinking / year)		-0.172***			
		(0.053)			
Tobacco consumption					
Smoker yes / no			-0.943***		
			(0.274)		
Cigarettes / day			-0.034		
			(0.039)		
Body Mass Index				-0.064**	
				(0.026)	
Debt					
Credit card debt (in €1,000)					-0.536***
					(0.088)
Other short-term debt yes / no					0.071
					(0.361)
Controls	-	Yes	Yes	Yes	Yes
Observations	-	403	403	398	351

Notes: Method: OLS. Unit of observation: subject. Dependent variables: *Support for policy proposal* is the support expressed for various policies, coded -2 (strong opposition) to 2 (strong support), averaged over policies for "All." A "neutral" response was possible only for the question about short-term lending. Because we asked subjects about *loosening* restrictions on short-term lending, we reverse-coded these responses for easier comparability (so that higher values correspond to greater support for tightening restrictions). *Beliefs about welfare effects* is the belief about the welfare effect of various policies, coded -2 (significantly worse off) to 2 (significantly better off), averaged over policies for "All." Controls: *Average mandate imposed on Chooser* is the maximum amount the Choice Architect allows the Chooser receive immediately (as in Section 3.3), averaged over rounds involving menus 3 and 4 in the Main condition. *Average belief exog. smaller choice set better for Chooser* (as in Section 3.2) encodes whether the Choice Architect considers the smaller opportunity set better, equally good, or worse for the Chooser than the unrestricted set, coded as 1, 0, and -1, respectively, averaged across the four rounds of the Exogenous Restriction Condition. All regressions control for session and order fixed effects. For variables measured with interval precision, we use midpoints for analysis. Samples: includes 303 subjects who participated in the main experiment, plus 100 subjects who participated in the Choice Distribution Information Condition (see Appendix B.4). Regressions in Panel C exclude subjects who chose not to disclose the personal characteristic of interest. Standard errors: clustered by subject. $^{*}p < 0.1$,$^{**}p < 0.05$,$^{***}p < 0.01$.

B.12 Beliefs about reasons for impatient choice

Here, we examine subjects' responses to non-incentivized survey questions regarding the reasons for choosing impatiently. We list six possible reasons a Chooser may select €4 today and €0 in half a year over €0 today and €10 in half a year. Choice Architects indicate agreement or disagreement on a scale of -2 to 2. Table A11 lists the reasons provided along with the corresponding mean responses. Choice Architects believe that plausible explanations include impatient predispositions and liquidity constraints, with significantly positive mean scores of 0.640 and 0.782, respectively. On average, Choice Architects consider inattention and random choice to be relatively unlikely explanations, and they judge the inability to implement one's objectives as downright implausible. Neither do Choice Architects believe that impatient choices are due to experiment-specific factors such as trust in the experimenter or technical issues with the monetary transaction.

Table A11: Non-incentivized assessment of the plausibility of particular reasons for impatience.

Reason	Mean plausibility rating
Objectives	
They are generally rather impatient.	0.640***
	(0.056)
They have an urgent need for money to pay for things like food or rent.	0.782***
	(0.069)
Implementation	
They did not pay attention and chose randomly.	-0.436***
	(0.059)
They meant to choose €0 today, €10 in half a year, but chose something else because some irrelevant event prevented them from choosing what they actually meant to choose (e.g. their hand trembled, or they confused the order of the radio-buttons).	-1.360***
	(0.053)
Experiment-related factors	
They are not sure whether the experimenter will really pay them half a year from now.	-0.838***
	(0.064)
They are not sure whether they will be able to receive the money in half a year from now, for instance because they no longer have a PayPal account.	-0.667***
	(0.062)

Notes Subjects were asked to suppose that a future experiment participant chooses €4 today and €0 in half a year over €0 today and €10 in half a year, and were asked to indicate how likely they think that each of the reasons listed motivated the participant's choice. Responses are given on the scale *extremely unlikely, unlikely, possible, likely, extremely likely*, encoded as -2, -1, 0, 1, 2, respectively. Standard errors in parentheses.

B.13 Performance on memory check

At the end of the experiment, Choice Architects reproduced their own choices from two of the twelve multiple decision lists about intertemporal choice they completed in the online component (in addition, they reproduced their choices from two lists of lotteries). The first of these lists used early and late payment amounts of €2 and €10, respectively, while the second list used €5 and €10. A total of 46.9% of subjects

correctly reproduced their responses for both decision lists. This number is potentially inflated due to the fact that 60.1% of all subjects chose the delayed option on each line of each of the two lists, which simplifies memorization and reconstruction.

C Vignette experiment

C.1 Design

To evaluate the possibility that differential beliefs about the effectiveness of various interventions contribute to our results in Section 5, we conduct a vignette study on Amazon Mechanical Turk. Its structure is similar to the part of our main experiment concerning policy judgments. Appendix D.5 displays the full text of the vignette study.

To control beliefs about the effectiveness of the interventions, we describe a specific individual, ask subjects to assume that the policy will affect that individual in a precisely specified fashion, and ask whether they would support the policy if it affected everyone the same way. In addition, we elicit beliefs about the effectiveness of the policies on potentially heterogenous members of the general population, permitting a test of whether beliefs about effectiveness are related to respondent characteristics in a way that would artificially generate ideals-projective paternalism.

We use four policies, displayed to subjects in random order: Alcohol taxes, sugary drinks taxes, retirement savings mandates, and restrictions on short-term, high-interest lending.[4] For each policy we describe an individual living in an Anglo-Saxon country other than the US, so respondents' attitude to paying taxes themselves should not influence their responses. Moreover, we describe all taxes as budget-neutral. Regarding sugary drinks taxes, for instance, we employ the following language: *"Anne, [is] a resident of Melbourne, Australia. She is 35 years old, 5 feet and 4 inches tall, and she weighs 190 pounds. ... For this question, assume that if the tax is introduced, Anne will reduce her consumption of sugary drinks so that her weight permanently drops to 145 pounds (from the previous 190 pounds). Given Annes height, this is a normal weight, according to the World Health Organization. If all residents of Melbourne were exactly like Anne, would you support or oppose the introduction of the tax?"*

After subjects provide their judgments, we also elicit respondents' own characteristics that allow for tests of mistakes-projective or ideals-projective paternalism for each of the policies. Specifically, we elicit weekly alcohol consumption, yearly binge drinking frequency, weight and height (to calculate BMI), a subjective assessment of the respondents' body-shape (underweight, healthy, overweight, etc.), the amount of the respondent's credit card debt, whether the subject has ever taken a payday loan, as well as the current stock of retirement savings and the respondent's current retirement savings rate.[5]

C.2 Analysis

We conducted the survey on the morning of February 1, 2019, with a total of 250 mTurk workers. Subjects received \$3 plus \$0.25 for each of eight attention check questions they answered correctly (two per policy). We retain the 161 subjects who correctly answered all attention check questions.

We define our independent variable by combining a subjects' responses to the two questions relating to any given policy and extracting the first principal component. We encode the resulting variable such

[4] We do not elicit attitudes to tobacco taxes because we conduct our survey with US subjects. The fraction of smokers in the US is substantially smaller than amongst our German laboratory subjects.

[5] We elicit additional characteristics, see Appendix D.5.

that negative coefficient estimates correspond to ideals-projective paternalism whereas positive coefficient estimates correspond to mistakes-projective paternalism.

For each policy, we perform an ordered probit regression of the support the respondent expresses for the policy under the assumption that all affected individuals are exactly the same as the person we described in the vignette.

Panel A of Table A12 displays the results. We find significant ideals-projective paternalism for alcohol taxes and for the retirement savings mandate. For sugary drinks taxes and for limits on short-term, high-interest lending, our coefficient estimates are not significantly different from zero. In no case are our estimates consistent with mistakes-projective paternalism.

Panel B uses as the dependent variable the support respondents express for the policy overall, when we ask them to consider that people are actually heterogeneous and the policy may create externalities. We find a significant negative relation between support and own characteristics for alcohol taxes and for the retirement savings mandate, and a marginally significant negative relation for regulation of payday lending. We continue to find a null effect for sugary drinks taxes.

The similarity of the results in Panels A and B, including for sugary drinks taxes, suggests that the relationship between policy support and the respondents' own characteristics is not contaminated by correlations between own characteristics and views of the policies' efficacy. We demonstrate this point more directly in Panel C, which uses beliefs about the effectiveness of the policies as a dependent variable. There is no significant association between respondents' own characteristics and their beliefs about the effectiveness of any of the three policies for which we elicited these beliefs.

Overall we conclude that our evidence of ideals-projective paternalism regarding real-world paternalistic policies is not contaminated by beliefs about efficacy systematically varying with respondent characteristics. Moreover, while our results show that ideals-projective paternalism is present for some but not all paternalistic policies among our US subjects, we find no indication of mistakes-projective paternalism.

Table A12: Results of the vignette study.

Policy	(1) Alcohol	(2) Sugar	(3) Retirement	(4) Loans
A. Policy support regarding person described in vignette				
Own characteristics	-0.226**	0.031	-0.282***	0.029
	(0.095)	(0.081)	(0.101)	(0.070)
oProbit thresholds				
1	-0.464	-1.922***	-1.654**	-0.822
	(0.691)	(0.736)	(0.716)	(0.666)
2	0.000	-1.458**	-1.123	0.402
	(0.694)	(0.722)	(0.708)	(0.653)
3	0.673	-0.933	0.097	1.338**
	(0.697)	(0.716)	(0.695)	(0.658)
Observations	161	161	146	158
B. Overall policy support				
Own characteristics	-0.269***	0.072	-0.273***	-0.148*
	(0.087)	(0.082)	(0.104)	(0.084)
oProbit thresholds				
1	-0.341	-1.982***	-2.189***	-1.131*
	(0.702)	(0.719)	(0.759)	(0.656)
2	0.164	-1.534**	-1.511**	-0.041
	(0.709)	(0.711)	(0.740)	(0.654)
3	0.913	-0.882	-0.500	0.926
	(0.712)	(0.701)	(0.721)	(0.659)
Observations	161	161	146	158
C. Effectiveness beliefs				
Own characteristics	-0.029	-0.055	-	-0.023
	(0.068)	(0.063)		(0.089)
Observations	161	161	-	158

Notes: Each column in each panel is a separate ordered probit regression that controls for the stage at which each policy was displayed, whether arguments in favor of or opposed to the interventions were mentioned first, the logarithm of the duration the subjects took to complete the sample, age, gender, and a dummy for each level of education. We did not elicit beliefs about the effectiveness of the retirement savings mandate. Standard errors clustered by subjects.

D Experiment design and implementation details

D.1 Experiment design: comprehensive description

The experiment consists of an online component and a laboratory component. Each encompasses multiple stages. Table A13 provides an overview, with stages listed in the order they were presented to subjects. Appendix sections D.3 and D.4 present verbatim translations of all instructions and questions, which we presented in German.

Table A13: Stages of the experiment.

Online component
1. Decision lists on intertemporal and risky choice
2. Non-incentivized questions
Laboratory component
1. Constructing Chooser's opportunity sets
2. *Additional decisions*
2.1 Surrogate choices
2.2 Attention test
2.3 Belief elicitation
2.4 Choice for oneself
2.5 Altruism / spite elicitation
3. *Real-world policies survey*
3.1 Policy judgments
3.2 Questions about reasons for impatient choice
3.3 Elicitation of characteristics related to real-world policies
4. Memory test

Online component The main part of the online component presents the six decision lists to elicit Choice Architects' own intertemporal choices, as described in Section 2.1. These are intermingled with six decision lists concerning risk taking. In these lists, subjects decide between winning amount y with probability p and 0 otherwise, or a sure amount $z \in \{1, 2, 3, 4, 5, 6\}$. We use $(y, p) \in \{(10, 0.5), (6, 0.7), (8, 0.5), (5, 0.8), (20, 0.2), (13, 0.3)\}$. The online component concludes with a battery of non-incentivized questions concerning age, gender, number of credit cards, and monthly expenses. Next, subjects select one of the four self-descriptive statements in panel B of Table 5, and make a selection from four similar statements concerning risk preferences. Subjects then provide answers to the non-numerical questions about risk and time preferences described in Falk et al. (2016). Finally, subjects provide information about the faculty at which their main course of study is offered, about the Bundesland in which they obtained their Abitur,[6] about their GPA in the Abitur exam, about their Abitur grade in mathematics, about their Abitur grade in German language, as well as about whether they had enrolled in honors classes in mathematics and / or German.

Laboratory component After reading general instructions about the experiment and passing a comprehension check, subjects make decisions in the following stages.

[6] The *Abitur* is the German university entrance qualification.

1. Constructing opportunity sets. Subjects read instructions about this stage of the experiment and complete two comprehension checks. Subsequently, in each of eleven rounds, Choice Architects first construct an opportunity set for the Chooser, and then reveal their beliefs about the extent to which receiving a specific subset rather than the full menu affects the Chooser's well-being, as described in Section 2.1.

2.1 Surrogate choices. In each of eleven rounds, Choice Architects make surrogate choices, and subsequently reveal their beliefs about the extent to which receiving the chosen option rather than the full menu affects the Chooser's well-being. The rounds correspond to the six menus used for intertemporal decisions, and the two menus used for the Induced Chooser Preferences condition. Subjects make surrogate choices concerning menu 5 four times, once for each of the Chooser self-descriptions in Table 5. Subjects do not see separate instructions for this part. Instead, in each round, subjects read "in this round you can make exactly one option available to the future experiment participant." The options are shown in a list from which the subject selects one. This part also removes the possibility of providing advise to the Chooser by the click of a button.

2.2 Attention test. Subjects answer the eight attention test questions reproduced and analyzed in section B.3. The instructions at the beginning of the online component inform subjects that their performance could completely determine their payment from the study. The instructions do not reveal any information about the test's content or focus, except that it will be about "what happened during this part."

2.3 Belief elicitation. Subjects read instructions concerning this stage and complete a comprehension check. In each of 11 rounds, they then reveal their beliefs about the distribution of the choices ten previous experiment participants have made when allowed to choose from the menus without restrictions, as described in Section 2.1. These are the same 11 rounds as in stage 2.1.

2.4 Choice for oneself. Subjects proceed through 8 rounds in which they select an option for themselves from each of the six menus in Table 1. Either menu 1 or menu 2 is subject to front-end delay (the same menu for which the choice set construction decision was subject to front-end delay). Subjects also make choices for themselves for each of the two rounds of the Induced Chooser Preferences condition. For these rounds, each subject is assigned one of the induced exchange rates $\hat{\delta}_i \in \{0, 0.5, 1\}$, which applies to both rounds.

2.5 Altruism / spite. Subjects decide whether to costlessly increase another Chooser's payment by €1, leave it unchanged, or decrease it by €1.

3.1 Policy judgments. Subjects rate four policy proposals concerning taxes on sugary drinks, alcohol, and tobacco, as well as about restrictions on short-term, high-interest loans. The tax policies concern Switzerland. We ask subjects to assume that the tax policies would be budget neutral. For each policy, we elicit the extent to which the subject supports or opposes its implementation in Switzerland. We also elicit beliefs about how the policy would change the welfare of the average citizen. The question about the welfare effects of alcohol taxes concerns adolescents and young adults rather than the average citizen, but is otherwise identical.

3.2 Questions about reasons for impatient choice and about reasons for restricting choice. Subjects first answer four questions about whether Choosers may have difficulties making (good) choices.[7] Second, subjects

[7]The questions are the following. Suppose an experiment participant can choose between the following options: Option A: Get €0 today, €10 in half a year, Option B: Get €4 today, €0 in half a year. 1. How easy do you think it is for the future experiment participant to know which of these two options is best for him? 2. How many out of 10 participants will choose the option that is really best for them? 3. How many out of 10 experiment participants will choose an option other than the one they actually wanted to choose? Suppose an experiment participant chooses "Option A: €0 today, €10 in 6 months from today". Given the experiment

provide judgments about the plausibility of six potential reasons for impatient choice listed and analyzed in Section B.12. Third, in each of five rounds, subjects are shown their own decisions from the choice set construction stage and are asked an open-ended question as to why they constructed the choice set the way they did. Four of these rounds correspond to the four Chooser types of Table 5 (subjects are shown the corresponding Chooser statement). The remaining round corresponds to a decision Choice Architects have made for a Chooser who would be selected at random.

3.3 Elicitation of characteristics related to real-world policies. Subjects provide information on their body mass index, their average alcohol consumption, their frequency of binge drinking, their smoking status and cigarette consumption, as well as whether they have ever taken out a short-term, high-interest loan, and if so, how quickly they repaid it. Subjects also rate their own political orientation on a left-right spectrum.

4. Memory test. Subjects are asked to reproduce their answers for four of the multiple-decision lists from the online component. Subjects learn of this test immediately before it begins. They receive €2 for exact replication. See section B.13 for questions and performance.

Order Half of Choice Architects, selected at random, proceed through the laboratory component in the order listed above, while the other half complete Stage 2.3 first.[8] Within each of these stages, the order in which subjects proceed through the conditions is individually randomized. Questions regarding real world policies, in contrast, are presented in the following order for all subjects: 1. sugary drinks taxes, 2. alcohol taxes, 3. tobacco taxes, 4. restrictions on short-term, high-interest lending.

Implementation All instructions are displayed on-screen. We intersperse them with four comprehension checks which subjects must pass in order to continue with the study, concerning (i) the incentive scheme and Induced Chooser Preferences condition, (ii) the manner in which Choosers rank and obtain outcomes, (iii) choices concerning Choosers in Stages 1 and 2.1, and (iv) the belief elicitation in Stage 2.3. Each comprehension check consists of at least five statements with at least two answer options each. The subject must select the correct response to each statement. In case of a mistake, the subject only learns that one or more of her responses were incorrect, but not which ones. Accordingly, it is exceedingly unlikely that subjects pass the comprehension checks by luck or by trial and error. Subjects who do not pass are referred back to the instructions (all subjects eventually passed).

All incentive payments are processed through PayPal; all transaction fees were paid by the researchers. The invitation email informs subjects of this fact and asks them to open a PayPal account if they do not already possess one.[9] Subjects received individualized links to ensure they would continue with the correct survey (implemented in Qualtrics) even though they completed the online component on a different machine than in the laboratory component.

Prior to fielding our laboratory experiment, we conducted pilot experiments on Amazon Mechanical Turk. Data from each pilot are consistent with our conclusions concerning projective paternalism.

participant chooses this option, is this a good or bad choice for this experiment participant? 4. Suppose the experiment participant chooses "€4 today, €0 in 6 months from today". Given the experiment participant chooses this option, is this a good or bad choice for this experiment participant?

[8] Only Choice Architects who reveal beliefs before constructing opportunity sets are in the Choice Distribution Information treatment, because the value of belief elicitation after information provision is debatable.

[9] PayPal accounts are free.

D.2 Sessions

We ran the experiment in 16 sessions at the Cologne Economics Research Lab. All materials were presented in German. The same head research assistant was present in every session. In each session, two additional research assistants were present to help. After the main sessions were concluded, we conducted five sessions to implement the Choice Architects' decisions that affected a Chooser. Each Chooser was presented with a single menu. Payments were determined according to Choosers' rankings of the options and the decisions the matched Choice Architects had made for the corresponding menu. While we do not use any data from the Chooser sessions in our analysis, we report the distribution of Choosers' choices in Appendix Section B.5. Table A14 lists the details of each session. Dates and times reflect availability of the lab and the research assistant.

Table A14: Session details.

Choice Architect sessions		
Session	Date	Time
1	6/14/18	4 pm
2	6/15/18	4 pm
3	6/19/18	10 am
4	6/19/18	1 pm
5	6/19/18	3.30 pm
6	6/20/18	10 am
7	6/20/18	1 pm
8	6/20/18	3.30 pm
9	6/26/18	10 am
10	6/26/18	1 pm
11	6/26/18	3.30 pm
12	6/27/18	1 pm
13	6/27/18	3.30 pm
14	7/3/18	1 pm
15	7/3/18	3.30 pm
16	7/4/18	2 pm

Chooser sessions		
Session	Date	Time
1	7/10/18	2 pm
2	7/10/18	3.30 pm
3	7/11/18	2 pm
4	7/11/18	3.30 pm
5	10/1/18	2pm

D.3 Instructions for the online component

[Horizontal lines represent screen breaks.]

Technical Check

To test whether your computer can display the study correctly, please copy the following number into the field below.

[Number]

If you do not see the number above, make sure Javascript is enabled on your web browser. Alternatively, use a different web browser.

This is a research study by the University of Cologne and the University of Toronto.

Parts

This study has two parts.

Online Part
Finish this part until [Date] (by midnight). This will only take a few minutes.

If you do not complete the online part, you will not be allowed to participate in the laboratory part.

Laboratory Part
Please appear in the laboratory at the time mentioned in the invitation email. This part will take between 1 and 2 hours.

PAYMENT

The payment for your participation in this study consists of two parts:

Base payment €12,50

[For sessions 1 and 2, the base payment was 9.50]
(€4.50 for the appearance at the laboratory part, and €8 for the completion of the laboratory part)

Bonus payment between €0 and €15

You will receive your base payment as soon as you have completed the laboratory part of the study.

The exact amount of your bonus payment, as well as the time at which you will receive it, depends on your decisions, as well as on chance.

The bonus payment and €5 of the base payment will be transferred via PayPal. For this purpose, you will provide your email address for the referral to the laboratory staff.

If you finish the online part but do not appear in the laboratory, you are entitled to pick up €2 at Prof. Ockenfels's office on weekdays between 10 a.m. and 12 noon until one week after the laboratory part's date.

At this stage, the subject gives informed consent. If no consent is given, the study is terminated.

Instructions for the online part

IMPORTANT: What happens in the online part of this study has no influence on what will happen in the laboratory part of the study.

In this part of the study you will participate in 12 decision rounds. Then you will answer some questions about yourself.

In each round you will see a list of six decisions, for example as follows:

Option A1	◯	◯	Option B1
Option A2	◯	◯	Option B2
Option A3	◯	◯	Option B3
Option A4	◯	◯	Option B4
Option A5	◯	◯	Option B5
Option A6	◯	◯	Option B6

These options will be replaced by specific amounts of money.

Your task is to select the option you truly prefer on each line.

With 25% probability, your total bonus payment will be determined by a single decision in one such list. In this case you will receive exactly what you have chosen on the corresponding line. (Which list and which decision this will be will be randomly decided by the computer at the end of this study.)

You should make every decision as if it were the one that counts - because it might be.

Bonus payment for this part
There is a 25% chance that your bonus payment from this study will be determined exclusively by this part of the study. (With the remaining probability of 75% it will be determined by a decision you will make in the laboratory part.)

In this case, at the end of the whole study, the computer will randomly select one of the decision rounds, and one decision from that round. You will receive exactly what you have chosen in this decision.

So it is in your interest to make every decision as if it were the one that counts because it may be!

Payments that depend on chance
Some options you can choose depend on chance (such as "Get €X with p% probability").

If you choose such an option, and it is randomly selected to determine your bonus payment, the following will happen. At the end of the laboratory part of this study, the study director will provide you with a die. You will be able to throw it once. The resulting number determines your final payout.

Late Payments
Some options you may choose involve receiving a certain amount of money at a given future date. This date can be up to six months after the laboratory part.

On behalf of the Ockenfels research group, we guarantee that if you decide for such an option, we will transfer exactly the specified amount at the specified time by PayPal.

The online part of this study starts now.

Your decisions and answers **will not** affect what will happen in the laboratory part of this study.

Please click NEXT.

[Decisions regarding time. The parameters X and Y are listed in footnote 7 of the main text.]

Please select an option on each line according to your real preference.

I would rather have . . .

. . . €X on the day of the laboratory part	◯ ◯	. . . €Y 1 month after the laboratory part
. . . €X on the day of the laboratory part	◯ ◯	. . . €Y 2 months after the laboratory part
. . . €X on the day of the laboratory part	◯ ◯	. . . €Y 3 months after the laboratory part
. . . €X on the day of the laboratory part	◯ ◯	. . . €Y 4 months after the laboratory part
. . . €X on the day of the laboratory part	◯ ◯	. . . €Y 5 months after the laboratory part
. . . €X on the day of the laboratory part	◯ ◯	. . . €Y 6 months after the laboratory part

[Decisions regarding risk. The parameters X and Y are listed in Appendix section D.1.]

Please choose one option on each line, depending on what you really prefer (all payments for this round will be made on the day of the lab experiment)
I would rather have . . .

. . . €X with probability p% and €Y with probability $(1-p)$%	◯ ◯	. . . €6 with certainty
. . . €X with probability p% and €Y with probability $(1-p)$%	◯ ◯	. . . €5 with certainty
. . . €X with probability p% and €Y with probability $(1-p)$%	◯ ◯	. . . €4 with certainty
. . . €X with probability p% and €Y with probability $(1-p)$%	◯ ◯	. . . €3 with certainty
. . . €X with probability p% and €Y with probability $(1-p)$%	◯ ◯	. . . €2 with certainty
. . . €X with probability p% and €Y with probability $(1-p)$%	◯ ◯	. . . €1 with certainty

Questions about yourself.

To conclude the online part, we would like to ask you some questions about yourself.

Please answer truthfully.

Your answers will not affect your payment or what will happen in the laboratory part of this study.

What is your gender?
[male; female; not listed above (e.g. genderqueer); prefer not to say]

How old are you?
[18 - 90]

How many credit cards do you hold?
[1; 2; 3; 4; 5; 6; 7; 8; 9; 10; >10; prefer not to say]

How much credit card debt do you currently have (in total, in Euros, across all your credit cards)?
[; 0; 1–250; 251–500; 501–750; 751–1000; 1001–1500; 1501–2000; 2001–3000; 3001–5000; 5001–10000; >10000]

How much money do you spend on average per month (including rent, food, mobility, etc.)?
[€ 0 - € 50; € 50 - € 100; € 100 - € 150; € 150 - € 200; € 200 - € 250; € 250 - € 300; € 300 - € 350; € 350 - € 400; € 400 - € 450; € 450 - € 500; € 500 - € 600; € 600 - € 700; € 700 - € 800; € 800 - € 900; € 900 - € 1000; € 1000 - € 1250; € 1250 - € 1500; € 1500 - € 1750; € 1750 - € 2000; € 2000 - € 2500; € 2500 - € 3000; more than € 3000; I prefer not to say]

Please select the statement that describes you best *[Choices were displayed in individually randomized order.]*
[I am a patient person. I am happy with this. (I often forego things in the present with regard to the future.); I am a patient person. I often regret my decisions. (Perhaps too often, I forego things in the present with regard to the future.); I am an impatient person. I am happy with this. (I rarely forego things in the present with regard to the future.); I am an impatient person. I often regret my decisions. (Perhaps too rarely, I forego things in the present with regard to the future.)]

Please select the statement that describes you best
[I am a risk-taker. I am happy with that. (I like to take risks.); I am a risk-averse person. I am happy with that. (I try to avoid risks if possible.); I am a risk-taker. I often regret my decisions. (I might be taking too many risks.); I am a risk-averse person. I often regret my decisions. (I might be trying to avoid risks too much.)]

Please select the statement that describes you best: I often do without things so that I can afford more later.
[Absolutely not like me; Very little like me; Not really like me; Neutral; A little like me; Very similar to me; Absolutely like me]

Are you generally a person who takes risks or do you try to avoid them?
[I try extremely hard to avoid risks; I try pretty hard to avoid risks; I try a little bit to avoid risk; I don't dislike taking risks; I'm quite willing to take risks; I'm extremly willing to take risks]

Please select the statement that describes you best: I tend to put things off until later, although it would be better to do them right away.
[Absolutely not like me; Very little like me; Not really like me; Neutral; A little like me; Very similar to me; Absolutely like me]

At which faculty do you study?
[Faculty of Economics, Management and Social Science; Faculty of Law; Faculty of Medicine; Faculty of Philosophy; Faculty of Mathematics and Natural Sciences; Faculty of the Humanities; I am not a student]

Which state conferred your Abitur?
[Baden-Württemberg; Bayern; Berlin; Brandenburg; Bremen; Hamburg; Hesse; Mecklenburg-Vorpommern; Niedersaxen; Nordrhein-Westfalen; Rheinland-Pfalz; Saarland; Sachsen; Sachsen-Anhalt; Schleswig-Holstein; Thüringen; I received the International Baccalaureate; I do not have an Abitur; I prefer not to say]

What was your Grade Point Average in the Abitur?
[1.0, 1.1, 1.2, . . . , 3.9, 4.0; I do not have an Abitur; I do not remember; I prefer not to say]

What was your Abitur grade in Mathematics?
[15 points (1+), 14 points (1), 13 points (1-), 12 points (2+), 11 points (2), 10 points, (2-), . . . , 3 points (5+), 2 points (2), 1 point (2-), 0 points; I do not have an Abitur; I do not remember; I prefer not to say]

What was your Abitur grade in German?
[15 points (1+), 14 points (1), 13 points (1-), 12 points (2+), 11 points (2), 10 points, (2-), . . . , 3 points (5+), 2 points (2), 1 point (2-), 0 points; I do not have an Abitur; I do not remember; I prefer not to say]

Have you taken an honors class in Mathematics in high school (Leistungskurs im Abitur)?
[Yes; No; I do not have an Abitur]

Have you taken an honors class in German in high school (Leistungskurs im Abitur)?
[Yes; No; I do not have an Abitur]

This is the end of the online part of this study

Please arrive at the laboratory on time.

This is a study of individual decision making. Therefore, please do not discuss this study with other people.

Please close this browser window.

(If you leave this window open, the laboratory part of this study will not start for you.)

D.4 Instructions for the laboratory component

Laboratory Part

Please enter your personal experiment code to ensure that you are proceeding with the correct questionnaire.[10]

Please enter the password provided by the experiment staff to start the laboratory part of this study.[11]

Instructions

Please read carefully.
This study contains multiple comprehension tests.
For simplicity, this study uses male pronouns throughout. They refer to both genders.

Payment for this study

The laboratory component of this study consists of 3 parts. Your decisions influence not only your own pay, but also that of future experiment participants. The study ends with some opinion questions, and some questions about yourself.

Affecting your own payment
There is a 75% chance that the lab component of the study will determine your bonus payment. (With the remaining 25% probability, your bonus payment will be determined by the online component.)

In this case your payment will be determined by exactly one of the three parts of the lab component. At the end of the study, the computer will randomly select a part and a decision you made in that part. This decision will be the only one that determines your bonus payment.

So you should make every decision as if it's the one that counts - because it may be!

Affecting the payment of other experiment participants

A part of this study consists of decisions that affect a future experiment participant. You will be able to influence that person's decision options and bonus payment.

At the end of this study, the computer selects *exactly one* decision you have made in this part. With a 1 in 4 chance we will match you with a future experiment participant. Your decisions in this study will then affect that person exactly as you have determined.

The future experiment participant's options for the bonus payment are determined entirely by the single decision of yours that the computer has randomly selected to be carried out.

You are the only person who determines the options of this future experiment participant. The future experiment participant will not make decisions that affect others; all his decisions will only affect himself.

None of the options that may determine your own payment in this study have been influenced by anyone else.

IMPORTANT: There is NO DECEPTION in this study. We will conduct the partner studies with the future experiment participants within the next 30 days and your decisions will affect future experiment participants with exactly the stated probability.

Anything else would violate the Ethics Protocol (UT36180) under which this study is conducted.

Some rounds are about money that you or the other experiment participant can receive at different times. Other rounds are about gold and silver coins. We now explain these one after the other.

Money at different points in time

In every decision concerning money at different points in time, the following choice will be made available (both for the future experiment participant, and for you):

Standard option:
Receive a bonus of €0 today and €15 in 6 months.

Alternative options such as the following may also be available:
Receive a bonus of €X today and €Y in 6 months.

[10] Each participant received an individual number in the invitation email. They were asked to sit at the computer terminal corresponding to their number. That computer contained the individual-specific link to the qualtrics survey that the subject had started in the online part. This mechanism ensures that subjects continue with their own survey, without requiring the storage of any identifiable data within the qualtrics survey.

[11] Subjects could begin the laboratory part only after every subject was seated at the correct terminal. At that stage, the experiment staff supplied the password required to continue.

(X and Y will be replaced by concrete amounts)

On behalf of the Ockenfels research group, we guarantee that if an experiment participant chooses such an option, we will transfer exactly the specified amount at the specified time via PayPal.

Gold and silver tokens

Some decisions in this study concern gold and silver tokens. If an experiment participant is paid with gold and silver tokens, the following will happen.

We will exchange all tokens into Euros and transfer them to the PayPal account of the experiment participant on the day of the study.

Value of gold and silver tokens

Each **gold token** is worth exactly €**1**.

The value of the silver tokens varies for different experiment participants.

For **1/3** of the participants the value of a **silver token is** €**1**.
For another **1/3** it is €**0.5**.
And for the last **1/3** it is €**0**.

Before the future experiment participant makes a choice, he learns exactly how many Euros he will receive per silver token. If you make a choice regarding the tokens for your own bonus, you will also know exactly how many Euros you will receive per silver token.

However, you will **not** know how many euros the future experiment participant will receive per silver token. He could get €1,€0.5, or €0 per silver token. All you know is that for the future experiment participant each of these cases is equally likely.

To ensure that you have understood these elements of the study correctly, please click on all true statements (and only those).[12]

- ☐ For 1/3 of the experiment participants the value of a silver token is €1. For another 1/3 it is €0.5. And for the last 1/3 it is €0. I will know which of them the future experiment participant will get, but the future experiment participant will not know.
- ☐ For 1/3 of the experiment participants the value of a silver token is €1. For another 1/3 it is €0.5. And for the last 1/3 it is €0. The future experiment participant will know which one of these he will get before making a choice, but I don't know which one he will get.
- ☐ My bonus payment is determined by three randomly selected decisions, one from each part of the study.
- ☐ My bonus payment is determined by exactly one decision from one part of the study.
- ☐ The options of the future experiment participant who is assigned to me with a 1 in 4 chance are completely determined by a single randomly selected decision of mine.
- ☐ All the decisions I make in the study do not affect anyone else, but determine my own payment.

If you feel that you have understood the instructions but still cannot continue, please raise your hand.

The three parts of the laboratory component of this study begin now.

You will receive the instructions for each part just before the corresponding part begins.

Part 1 of 3

The decisions you make in this part will affect another experiment participant's bonus payment.

This is the longest part of the study, and will take about twice as long as the other two parts.

Please make all decisions in this section carefully.

How the future experiment participant is affected by your decisions in this part

This part has 25 rounds. Each round has two halves.

The computer randomly selects one round from this part and one of the two halves within this round. Every round and every half is equally likely.

The bonus payments and decision options of the future experiment participant will be determined by exactly this decision of yours.

So you should make every decision as if it were the decision that will affect the future experiment participant. Because it may be!

How this part affects your own bonus payment

It is in your own interest to be attentive. At the end of this part there will be a test of what happened during this part. It is possible that your bonus payment for this study will be determined entirely by your performance on this test.

We now explain what will happen in each of the two halves of each round.

[12] All statements were displayed in individually randomized order. A subject could continue only if he had correctly marked all of the statements. In case of an error, the participant did not receive feedback about which of the statements was marked incorrectly.

First half of each round

Available and unavailable choice options
The future experiment participant will be able to choose his bonus payment from a set of options as follows:[13]

Opportunity set X

☐	Option A
☐	Option B
☐	Option C

(Options A, B, and C will be replaced by concrete payments.)

You decide whether all options should be available to the future experiment participant or whether one or more of them will be unavailable. (You must make at least one option available in each round.)

We ask you to make these decisions with care. There are no right or wrong decisions. These decisions do not affect your own bonus payment.

The details by which you will make the choice options of the future experiment participant available or unavailable will vary across the rounds.

Why such decisions?

People have different views about influencing other people's choices and decisions.

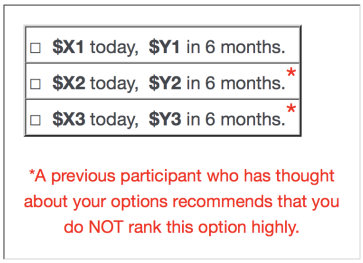
There are no right or wrong answers. We ask you to make exactly those decisions that reflect your genuine views.

Messages to the future experiment participant

In addition to making options available or unavailable, you will be able to send messages to future experiment participants if you wish. There are two types of messages in this study.

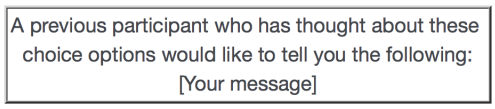
Click messages

In some rounds you will be able to click a button to advise the future experiment participant not to rank an option highly. If you do, the future experiment participant will see the following when making decisions about his options:



Free-form messages

If you wish, you can also send a freely formulated message to the future experiment participant. In this case, the future experiment participant will see the following when deciding between his options:



[This screen is only displayed to participants in the Choice Distribution Information treatment]

How previous experiment participants have chosen from the options

In addition to making options available or unavailable, in each round you can click a button to view a table like the following.

N1 of 100 previous experiment participants chose	€X1 today and €Y1 in six months
N2 of 100 previous experiment participants chose	€X2 today and €Y2 in six months
N3 of 100 previous experiment participants chose	€X3 today and €Y3 in six months

[13] There is no natural-sounding translation of the expression 'opportunity set' in German. We have used the semantically closest expression 'Warenkorb'. Its literal translation is 'basket of goods'.

This table displays how 100 previous experiment participants have chosen between the options when all options were available (X1, Y1, N1, etc. will be replaced by numbers).

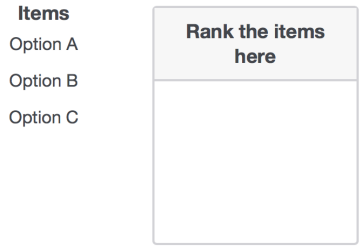
How the future experiment participant will choose

The future experiment participant will **not** see which of the options are available in an opportunity set and which are not.

Instead, he will see all options and rank them according to his preference, regardless of whether they are available or not. The participant then receives the available option that he has ranked the highest.

The experiment-participant will know that not all options may be available, and that he will receive the option he has ranked the highest among those options that are available. This is all he knows about how his payments are determined.

Here you can see how the future experiment participant can rank his options. We will ask him to put the option he most wants first and the option he least wants third (To try this, drag the options into the box on the right.)



Example

Suppose there are three options, A, B and C.

Also suppose you make option A unavailable for the future experiment participant.

For example, the future experiment participant might rank A first, then B, and C last. In this case, the experiment participant will receive B. The reason is that A is not available, and the participant ranks B higher than C.

To ensure that you understand how your decisions affect the other experiment participant, please answer the following questions.

Each question concerns three options, A, B and C.

Assume that the future experiment participant ranks B at the top, followed by C, and that he ranks A lowest.

If all options are available, which will the future experiment participant receive?
[Option A, Option B, Option C.]

If Option A is not available, which option will the future experiment participant receive?
[Option A, Option B, Option C.]

If Option B is not available, which option will the future experiment participant receive?
[Option A, Option B, Option C.]

If Option C is not available, which option will the future experiment participant receive?
[Option A, Option B, Option C.]

If Option B and C are not available, which option will the future experiment participant receive?
[Option A, Option B, Option C.]

Second Half of Each Round

In the second half of each round you will see two opportunity sets according to which the choices of future experiment participants could be determined, such as in the following example.

Opportunity Set Right	Opportunity Set Left
Option A	Option A
Option B	~~Option B~~
Option C	Option C

In this example, all options are available in the opportunity set on the left. Option B is not available in the opportunity set on the right.

Your decision is

- whether the future experiment participant should receive the option that he has ranked highest from the opportunity set on the left, or
- whether the future experiment participant should receive the option that he has ranked the highest from the opportunity set on the right; AND additionally the amount €Z should be added to / deducted from his base payment for the experiment.

You make several such decisions, for different values of Z, in a list such as this one:

The bonus payment of the future participant should be determined by ...			
... opportunity set Left and his base payment remains unchanged	◯	◯	... opportunity set Right and his base payment is raised by €0.5.
... opportunity set Left and his base payment remains unchanged	◯	◯	... opportunity set Right and his base payment is raised by €0.3.
... opportunity set Left and his base payment remains unchanged	◯	◯	... opportunity set Right and his base payment is raised by €0.2.
... opportunity set Left and his base payment remains unchanged	◯	◯	... opportunity set Right and his base payment is raised by €0.1.
...			...

Implementation

If this half of a round is randomly selected for the implementation, the following will happen:

The computer randomly selects one of the lines in the list. The decision you made in this line is carried out.

Therefore, you should make every decision on each line as if it were the one that counts. Because it may be!

In some rounds Opportunity Set Left and Opportunity Set Right may be the same. This is intentional. In these cases too, please make a careful choice in each line.

In the second half of each round we will also ask you:

In your opinion, which opportunity set is better for the future experiment participants' own good?

This question refers to **your opinion** about which opportunity set is better for the future experiment participant's own good – **regardless of whether the future experiment participant would agree with you or not!**

To ensure that you have understood these elements of the study correctly, please click on all true statements (and only those).[14]

- ☐ The computer will carry out all decisions I make in this part.
- ☐ I can make some options available and others not, if I think that is right.
- ☐ I can make all the options available if I think that is right.
- ☐ I can make all but one option unavailable if I think that's right.
- ☐ There are NO right or wrong decisions, I can make available or unavailable whatever I think is right (as long as at least one option is available for the future experiment participant).
- ☐ There ARE right or wrong decisions, I can NOT just make available or unavailable whatever I think is right.
- ☐ The future experiment participant will only see the options I make available and will not even see the options I make unavailable.
- ☐ The future experiment participant will rank all options. He will not know which ones are available or unavailable. His bonus payment is based on the option he has ranked most highly amongst those that are available.
- ☐ The computer carries out ONE randomly selected decision. This decision is equally likely from the first half of a round as from the second half of a round. And each round is equally likely.

If you feel that you have understood the instructions but still cannot continue, please raise your hand.

First half of round X

(Note: The subject is informed that they have been matched with another participant that's either impatient and unhappy, impatient and happy, patient and unhappy, patient and happy, or they are informed that they have been matched randomly.)

If this round is implemented, it will concern an experimental participant who says about himself:

I'm an impatient person. I often regret my decisions. (Perhaps too often, I forego things in the present with regard to the future.)

[14] All statements were displayed in individually randomized order. A subject could continue only if he had correctly marked all of the statements. In case of an error, the participant did not receive feedback about which of the statements was marked incorrectly.

Which of the choice options will be available to the future participant?
(You must make at least one option available)

	Available	Unavailable	Recommend against
€X1 today, **€Y1** in 6 months from today.	○	○	☐
€X2 today, **€Y2** in 6 months from today.	○	○	☐
€X3 today, **€Y3** in 6 months from today.	○	○	☐

If you have a message for the future participant, enter it here:

(Note: For surrogate choices, subjects are shown .)

In this round you can make exactly one option available to the future experiment participant. Which option will that be?
○ €**X1** today €**Y1** in 6 months from today
○ €**X2** today €**Y2** in 6 months from today
○ €**X3** today €**Y3** in 6 months from today

Second half of round X

(Note: The subject is informed that they have been matched with another participant that's either impatient and unhappy, impatient and happy, patient and unhappy, patient and happy, or they are informed that they have been matched randomly.)

If this round is implemented, it will concern an experimental participant who says about himself:

I'm an impatient person. I often regret my decisions. (Perhaps too often, I forego things in the present with regard to the future.)

Choice Set Left	Choice Set Right
€**X1** today, €**Y1** in 6 months	€**X1** today, €**Y1** in 6 months
€**X2** today, €**Y2** in 6 months	€**X2** today, €**Y2** in 6 months
€**X3** today, €**Y3** in 6 months	~~€**X3** today, €**Y3** in 6 months~~

Which choice set is better for the future participant's own good?

Choice set Left	Both equal	Choice set Right
○	○	○

The bonus payment of the future participant should be determined by ...			
... choice set Left, and his base payment remains unchanged	○	○	... choice set Right and his base payment is raised by €1.
... choice set Left, and his base payment remains unchanged	○	○	... choice set Right and his base payment is raised by €0.5.
... choice set Left, and his base payment remains unchanged	○	○	... choice set Right and his base payment is raised by €0.2.
... choice set Left, and his base payment remains unchanged	○	○	... choice set Right and his base payment is raised by €0.1.
... choice set Left, and his base payment remains unchanged	○	○	... choice set Right and his base payment is lowered by €0.1.
... choice set Left, and his base payment remains unchanged	○	○	... choice set Right and his base payment is lowered by €0.2.
... choice set Left, and his base payment remains unchanged	○	○	... choice set Right and his base payment is lowered by €0.5.
... choice set Left, and his base payment remains unchanged	○	○	... choice set Right and his base payment is lowered by €1.

[The subject proceeds through the remaining rounds of part 1 in the same fashion]

First half of round X

This round concerns a random future experiment participant.

In this round you can make exactly one option available to the future experiment participant. Which option will that be?
The bonus payment of the future experiment participant will be ...

○ €**X1** today €**Y1** in 6 months from today

◯ €**X2** today €**Y2** in 6 months from today
◯ €**X3** today €**Y3** in 6 months from today

If you have a message for the future participant, enter it here:

Second half of round X

This round concerns a random future experiment participant.

Choice Set Left	Choice Set Right
€**X1** today, €**Y1** in 6 months	€**X1** today, €**Y1** in 6 months
€**X2** today, €**Y2** in 6 months	~~€**X2** today, €**Y2** in 6 months~~
€**X3** today, €**Y3** in 6 months	~~€**X3** today, €**Y3** in 6 months~~

Which choice set is better for the future participant's own good?

Choice set Left	Both equal	Choice set Right
◯	◯	◯

The bonus payment of the future participant should be determined by . . .

. . . choice set Left, and his base payment remains unchanged	◯	◯ . . . choice set Right and his base payment is raised by €1.
. . . choice set Left, and his base payment remains unchanged	◯	◯ . . . choice set Right and his base payment is raised by €0.5.
. . . choice set Left, and his base payment remains unchanged	◯	◯ . . . choice set Right and his base payment is raised by €0.2.
. . . choice set Left, and his base payment remains unchanged	◯	◯ . . . choice set Right and his base payment is raised by €0.1.
. . . choice set Left, and his base payment remains unchanged	◯	◯ . . . choice set Right and his base payment is lowered by €0.1.
. . . choice set Left, and his base payment remains unchanged	◯	◯ . . . choice set Right and his base payment is lowered by €0.2.
. . . choice set Left, and his base payment remains unchanged	◯	◯ . . . choice set Right and his base payment is lowered by €0.5.
. . . choice set Left, and his base payment remains unchanged	◯	◯ . . . choice set Right and his base payment is lowered by €1.

[The subject proceeds through the remaining rounds of part 2 in the same fashion]

Please answer all questions about this part
If the computer selects this part to determine your payment, your bonus is calculated as follows: You receive €1 for each correct answer, and €0 for each incorrect answer.

In some rounds I could remove options, but I could never add options for the future experiment participant.
[True, False]

For all options which the future experiment participant could receive, the early payment will always be on the day of the experiment.
[True, False]

For some options that the future experiment participant could have received, the late payment will occur up to 7 months after the day of the study.
[True, False]

In some rounds, I had to make a single option available to the future experiment participant; I could not make multiple options available, even if I wanted to.
[True, False]

Some rounds concerned gold tokens and silver tokens. These rounds only concerned money that would be paid soon after the experiment, not money that would be paid only months after the experiment.
[True, False]

Other experiment participants could receive €0.40, €0.50, or €0.60 per silver token, each with the same probability.
[True, False]

In some lines of the decision lists, the base payment of the future participant could be increased. In other rounds it could be decreased.
[True, False]

If I made some options unavailable, this means that the other experiment participant cannot see that option, and therefore does not need to think about these options.
[True, False]

Part 2 of 3

This part has 12 rounds. Each round follows the same structure.

Previous experiment participants decided between different options from shopping baskets of three options each as follows:

Shopping basket X

☐	Option A
☐	Option B
☐	Option C

(Options A, B, and C were concrete payments.)

In each round we show you a different shopping basket. Your task is to estimate how previous participants chose between the options in the given shopping basket.

(In one of the 12 rounds we ask you instead to estimate which of 4 descriptions each of the previous participants considered the best description of themselves.)

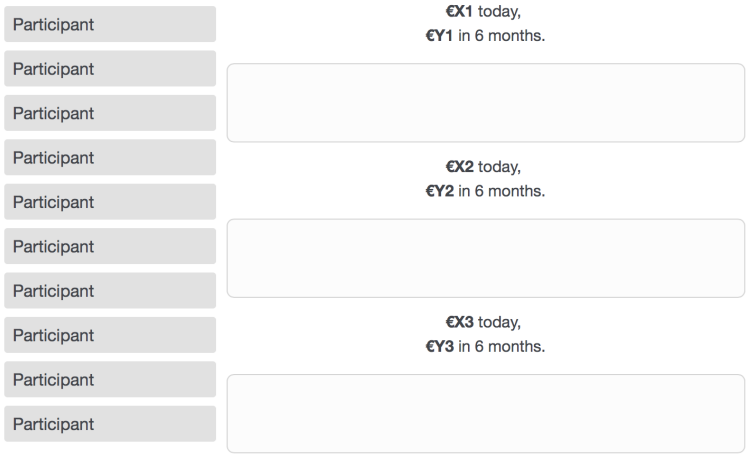
You will make these estimates as follows:

On the left side of the graph below you see "Participant" written 10 times. Each represents one of the 10 experiment participants. You can move them using Drag & Drop.

On the right side you see three fields. Each of them is a choice option.

Your task is to sort each of the 10 experiment participants on the left into one of the boxes on the right, depending on your estimate of how previous experiment participants actually chose from these options.

If you estimate that X of the previous participants had selected option A, that Y had selected option B, and that Z had selected option C, place X of the "Participants" in the "Option A" container, Y of the "Participants" in the "Option B" container, and Z of the "Participants" in the "Option C" container.



How your estimate will affect your bonus payment

Your payment from this study could be determined solely by this part of the study! (The computer will randomly decide whether you will be paid for this or another part.)

If so, the following will happen.

The computer randomly draws a round from this part.

We have data on how 10 previous experiment participants chose among these choice options. We compare your estimate with what the previous experiment participants actually chose.

If your estimate for the selected round is correct, your bonus payment will be €10, which you will receive today via PayPal.

If your estimate is incorrect compared to what previous experiment participants actually did, the following will happen.

Suppose you have assigned too many experiment participants to one option and too few to another option (compared to what the previous experiment participants actually chose).

We then take one of the fields "experiment participant" from one of the containers that has too many fields "experiment participant", and place it in one of the containers that has too few. We will do this until there are exactly as many "experiment participants" fields in each container as we actually observed with the 10 previous experiment participants.

For each "experiment participant" field that we put in a different container in this way, we will deduct €1 from the €10 that you would receive if you had perfectly estimated the decisions of the previous experiment participants. The rest you will receive as a bonus payment.

Therefore, you can expect that you will earn the most with this study if you think carefully about how the previous experiment participants actually decided and place the experiment participants in the containers accordingly.

Gold and silver coins

Some rounds involve decisions on gold and silver coins.

Their estimate concerns former **experiment participants** who knew whether they would receive €1, €0.5, or €0 per silver coin **before** making their decisions.

Please select all true statements. You can continue as soon as you have classified all statements correctly.

- ☐ I will make the most money in this part if I put all the experiment participants in the same container.
- ☐ I will make the most money in this part if I put the same number of experiment participants in each container.
- ☐ I will make the most money in this part by placing the experiment participants in the containers, according to my best estimate of how the previous experiment participants actually decided.
- ☐ If I put too many experiment participants in one container and too few in another (compared to what the previous experiment participants actually chose), I lose €1 for each experiment participant that I put in the wrong container.
- ☐ My answers in this section do not affect my payment.
- ☐ My payment from this study could be determined solely by this part of the study!

If you feel that you have understood the instructions but still cannot continue, raise your hand.

Part 3 of 3

In this section you make decisions that only affect your own payment.

This part has 8 rounds. You will make a selection in each round. At the end of the experiment, the computer randomly selects **exactly one** of your options. If this part determines your payment for this study, this option is the only one that counts.

So you should make every decision as if it were the one that counts. Because it maybe be!

[In rounds corresponding to the Induced Chooser Preferences condition, subjects see the following screen.]

Please choose one of the following options, depending on what you actually prefer.

- ◯ X1 gold coins, Y1 silver coins
- ◯ X2 gold coins, Y2 silver coins
- ◯ X3 gold coins, Y3 silver coins

Value of gold and silver coins
Each gold coin is worth €1.
The value of the silver coins differs between the experiment participants. For you every silver coin is worth €0.5.
This payment will be made today.

[In rounds except those corresponding to the Induced Chooser Preferences condition, subjects see the following screen. In the round with front-end delay, the dates are '1 week from today' and '6 months plus a week from today']

Please choose one of the following options, depending on what you actually prefer.

- ◯ € X1 today, € Y1 in 6 months
- ◯ € X2 today, € Y2 in 6 months
- ◯ € X3 today, € Y3 in 6 months

Completion payment for future experiment participants

There is a 1 in 4 chance that you will be assigned to a second future experiment participant. This is another experiment participant than the one for which you have decided what choice options they should get.

Like yourself, this future experiment participant will receive a base payment of € 9.50 for the completion of the study.

However, you can decide to change his base payment. The base payment of the future experiment participant will be exactly what you decide.

What completion payment should the second future experiment participant receive?

(This is a person other than the one whose options you have made available or unavailable)

- ◯ The future experiment participant should receive a base payment of € 10.50 instead of € 9.50.
- ◯ The future experiment participant should receive the planned base payment of € 9.50.
- ◯ The future experiment participant should receive a base payment of € 8.50 instead of € 9.50.

Questions about your opinion and about yourself

The last part of this study consists of a questionnaire. Please answer the questions honestly. Your answers do not affect your payment or the payment of other people from this study.

We would first like to ask your opinion on four policy proposals.

The first three proposals concern Switzerland. That country is very similar to Germany in many aspects. Since Switzerland has only one tenth of Germany's population and is not a member of the European Union, policy changes in Switzerland have no direct effect on Germany.

Proposal 1: Taxes on high-sugar beverages

Several countries around the world levy taxes on beverages with a high sugar content (which is associated with obesity). These countries include Hungary, Ireland, Norway, the Philippines, the United Arab Emirates, Great Britain and others.

There are many overweight people in Switzerland, as in the vast majority of developed countries.

What is your attitude towards a tax that would increase the price of sugary beverages in Switzerland by 20% (income tax would be reduced so that the government would earn the same tax revenue as before)?

Switzerland should . . .
[definitely not introduce such a tax, probably not introduce such a tax, probably introduce such a tax, definitely introduce such a tax.]

What do you think the effect of such a tax would be?

If such a tax were introduced, people in Switzerland would be on average. . .
[significantly worse off, a little worse off, neither better nor worse off, a little better off, significantly better off.]

Proposal 2: Taxes on alcoholic beverages

Binge drinking is the excessive consumption of alcoholic beverages with the aim of getting heavily drunk. Binge drinking is sometimes considered problematic. One reason is that binge drinking might be harmful to health.

It is therefore being discussed whether alcohol taxes should be increased in order to make binge drinking more expensive and correspondingly less frequent. Specifically, the price of cheap alcohol (spirits and cheap wines) could be increased disproportionately in percentage terms by charging the tax per liter of pure alcohol in the beverage.

What is your attitude towards a tax that would increase the price of spirits and cheap wines in Switzerland by 50%, on average (income tax would be reduced so that the government would earn the same tax revenue as before)?

Switzerland should ...
[definitely not introduce such a tax, probably not introduce such a tax, probably introduce such a tax, definitely introduce such a tax.]

What do you think the effect of such a tax would be?
If such a tax were introduced, adolescents or young adults in Switzerland would be on average...
[significantly worse off, a little worse off, neither better nor worse off, a little better off, significantly better off.]

Proposal 3: Taxes on tobacco

The harmful effects on health and the addictive potential of smoking cigarettes have been proven scientifically.

There is therefore a discussion as to whether taxes on cigarettes and other tobacco products in Switzerland should be further increased in order to reduce consumption and to deter young people more from smoking.

What is your attitude towards a tax that would increase the price of cigarettes of other tobacco products in Switzerland by an average of half the current price (income tax would be reduced so that the government would earn the same tax revenue as before)?

Switzerland should ...
[definitely not introduce such a tax, probably not introduce such a tax, probably introduce such a tax, definitely introduce such a tax.]
What do you think the effect of such a tax would be?

If such a tax were introduced, people in Switzerland would be on average...
[significantly worse off, a little worse off, neither better nor worse off, a little better off, significantly better off.]

Proposal 4: Restrictions on short-term loans

Short-term loans have been available in Germany since 2010. Such loans typically finance consumption, ranging from €50 to €3000, and must be repaid after 30 or 60 days.

The interest that can be charged on such loans is high, but limited by law. Therefore, people with very bad credit cannot obtain such loans, even if banks would be willing to grant such loans at very high interest rates and borrowers with very low credit ratings would be willing to pay very high interest rates for them.

It is being discussed whether the restrictions should be loosened. Advocates argue that many people who really need the money will otherwise not get credit. Opponents argue that such loans are very expensive and people can end up debt cycles.

How do you feel about that? The market for short-term loans in Germany should...
[be severely restricted (in this case, far fewer people can receive a short-term loan), be somewhat restricted (in this case, fewer people can get a short-term loan), remain unchanged, be somewhat liberalized (in this case, more people can get a short-term loan), be severely liberalized (in this case, far more people can get a short-term loan)]
What do you think the effect of such liberalization would be?

If this market were strongly liberalized, the average German would be...
[significantly worse off, a little worse off, neither better nor worse off, a little better off, significantly better off.]

We would now like to ask you some questions about the experimental decisions

The questions on this page concern the following scenario:

Suppose an experiment participant can choose between the following options:

- Option A: Get €0 today, €10 in half a year
- Option B: Get €4 today, €0 in half a year

How easy do you think it is for the future experiment participant to know which of these two options is best for him?
[Very easy, Rather simple, rather difficult, very difficult]

How many out of 10 participants will choose the option that is really best for them?
[0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10]

How many out of 10 experiment participants will choose an option other than the one they actually wanted to choose?
[0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10]

Suppose an experiment participant chooses "Option A: €0 today, €10 in 6 months from today". Given the experiment participant chooses this option, is this a good or bad choice for this experiment participant?
[Most likely bad, Very likely bad, Somewhat likely bad, Somewhat likely good, Very likely good, Most likely good]

Suppose the experiment participant chooses "€4 today, €0 in 6 months from today". Given the experiment participant chooses this option, is this a good or bad choice for this experiment participant?
[Most likely bad, Very likely bad, Somewhat likely bad, Somewhat likely good, Very likely good, Most likely good]

Suppose an experiment participant can choose between the following options:

- Option A: €0 today, €10 in 6 months from today, and
- Option B: €4 today, €0 in 6 months from today

Assume the experiment participant chooses the less patient option "€4 today, €0 in 6 months from today".

How likely do you think the following reasons are for someone making such a decision?

	Very unlikely	Unlikely	Maybe	Likely	Very likely
They are generally impatient.	○	○	○	○	○
They are not sure if they can receive the money in half a year, for example, because they will no longer have a PayPal account.	○	○	○	○	○
They actually wanted to choose "Get €0 today, €10 in half a year," but chose the other option because an irrelevant event prevented them from choosing what they actually wanted (e.g. their hand trembled, or they confused the order of radio buttons).	○	○	○	○	○
Today they urgently need money to pay expenses such as food or rent.	○	○	○	○	○
They are not sure if the study director would really pay them in half a year.	○	○	○	○	○
They were not attentive and chose randomly.	○	○	○	○	○

In a previous part of this study, you decided to make certain options available or unavailable.

We would like to ask you to explain in your own words: Why did you make the decisions you did?

Specifically we ask you about your decisions in the following round:

The round was about an experimental participant who said about himself:

I'm an impatient person. I often regret my decisions. (I probably do without the present too seldom with regard to the future.)

You have made the following options available: ...

You have made the following options unavailable: ...

Your answer:

[The subject answers this question five times, once for each of the four Chooser types of Table 5, and once for a randomly selected Chooser.]

Finally, please answer the following questions about yourself truthfully

How many alcoholic beverages did you consume on average per week, calculated over the last 12 months? (1 alcoholic drink = 0.2 liter beer, 0.1 liter wine, 1 shot of spirits or liquor)
[10 or more per week, 5-10 per week, 3-5 per week, 1-3 per week, less than 1 per week but some, none]

How often did you have 4 or more alcoholic beverages within a 2-hour period over the last 12 months?
[7 days per week, 5-6 days per week, 3-4 days per week, 2 days per week, 1 day per week, 2-3 days per month, 3-11 days over the last year, 1-2 days over the last year, never]

Sometimes people need to borrow money quickly and take out short-term, high-interest loans. Such loans include consumer loans of EUR 5000 or less with less than 1-year maturity, payday loans, pawn shop loans, or rent-to-own loans (but not credit card debt).

Do you have experience with short-term, high-interest loans?
[No, I've never taken out a loan like this before; Yes, I have taken such a loan in the past, but I don't do it regularly; Yes, I regularly take out such loans]

If you have taken out a short-term, high-interest loan (for example, payday loan or pawnbroking) in the past:
How did the repayment work? (If you have taken out several such loans or take them out regularly, please indicate what typically happens.)

[I repaid it in full on my next payday without taking out another loan immediately; I repaid it in full or in part on my next payday, but took up another loan to pay it; I held the loan longer than originally planned, but later repaid it without immediately taking out another loan; I held the loan longer than originally planned, but later paid it back and took out another loan to pay the other one; I was unable to pay it back and I went bankrupt with the loan; I've never taken out a loan like this before]

Which of the following categories describes you best? I am...
[severely underweight (BMI <16); underweight (BMI between 16 and 18.5); normal weight, on the lighter side (BMI between 18.5 and 21.7); normal weight, on the heavier side (BMI between 21.7 and 24.9); overweight (BMI between 24.9 and 29.9); severely overweight (BMI > 29.9); I prefer not to say]

[Subjects can click a button labelled 'calculate BMI'. If they do so, a window pops up prompting subjects to enter their weight and height. Upon clicking enter, the subjects can then view their BMI.]

Are you a smoker, or have you ever been a smoker? *[I am a non-smoker, and have never been a smoker; I now am a non-smoker, but I used to be a smoker; I am an occasional smoker; I am a smoker]*

How many cigarettes do you consume per day, based on the last 12 months (excluding e-cigarettes)?
[40 (2 packs) or more per day; Between 20 (1 pack) and 40 (2 packs) per day; Between 10 (1/2 pack) and 20 (1 pack) per day; Between 1 and 10 (1/2 pack) per day; I am a non-smoker]

What is your political orientation?
[left, center-left, left-of-center, centrist, right-of-center, center-right, right]

Memory test

As the final part of this study, you have the opportunity to earn an additional €2. These will be added to all your other payments via PayPal.

In the online part of this study you made a number of decisions about your bonus payment.

In this section we ask you to reproduce these decisions if you remember them.

It is **not** about what options you actually prefer now. It is about clicking on **the same** options as in the online part, even if you might prefer a different option now.

We will show you 4 lists of such decisions. If for all lists you click on the same options as in the online part of this study, we will add €2 to your PayPal payment today.

In each line, please select the SAME option you selected in the online part of this experiment (regardless of which option you currently prefer).

In the online part of this experiment I had chosen the following options:
I would rather have ...

...€X on the day of the laboratory part	◯ ◯	...€Y 1 month after the laboratory part
...€X on the day of the laboratory part	◯ ◯	...€Y 2 months after the laboratory part
...€X on the day of the laboratory part	◯ ◯	...€Y 3 months after the laboratory part
...€X on the day of the laboratory part	◯ ◯	...€Y 4 months after the laboratory part
...€X on the day of the laboratory part	◯ ◯	...€Y 5 months after the laboratory part
...€X on the day of the laboratory part	◯ ◯	...€Y 6 months after the laboratory part

---

In the online part of this experiment I had chosen the following options: I would rather have ...

...€X with probability p% and €Y with probability $(1-p)$%	◯ ◯	...€6 with certainty
...€X with probability p% and €Y with probability $(1-p)$%	◯ ◯	...€5 with certainty
...€X with probability p% and €Y with probability $(1-p)$%	◯ ◯	...€4 with certainty
...€X with probability p% and €Y with probability $(1-p)$%	◯ ◯	...€3 with certainty
...€X with probability p% and €Y with probability $(1-p)$%	◯ ◯	...€2 with certainty
...€X with probability p% and €Y with probability $(1-p)$%	◯ ◯	...€1 with certainty

---

This is the end of this study

Thank you for your participation!

Your Payment

You will receive your showup payment of €4.50 in cash, and your base payment of €5 via PayPal. You will also receive the bonus payment below.

The computer has randomly determined that your bonus payment for this study is determined by . . .

Your bonus payment therefore is €5

Your responses to the memory coincide with your choices in the online part. Therefore you will receive an additional €2.

Do you have any comments on this study?

Please click here if you wish to withdraw your data from this study. If you choose to do so, you will receive €1 for your participation in this study. You will not receive any other payment, and the decisions you have made in this experiment will not affect any future experiment participant.

☐ I want to withdraw my data from the study

Now, please raise your hand.

Someone from the experiment staff will take care of your payment.

D.5 Vignette experiment

Subjects were shown the four policies in individually randomized order. Subjects were paid $3 for completion, plus an additional $0.25 for each of eight attention check questions (two per policy) they answered correctly.

[Horizontal lines represent screen breaks.]
Policy 1 of 4: Sugary drinks taxes in Melbourne, Australia
The next questions are all about Anne, a resident of Melbourne, Australia. She is 35 years old, 5 feet and 4 inches tall, and she weighs 190 pounds. According to a classification by the World Health Organization, Anne is thus moderately obese. Sugary drinks such as Coca Cola are a major contributor to the obesity epidemic, according to a study by Harvard University. Currently, Australia does not have a sugary drinks tax. Suppose the city of Melbourne therefore considers introducing a sugary drinks tax. If passed, the price of all sugar-sweetened beverages would double. The value added tax on all other goods would be lowered, so that the state would collect the same overall amount of taxes as without the change.

With the sugary drinks tax, people would consume fewer sugar-sweetened beverages. On the one hand, this would lower their body weight. On the other hand, they would less frequently enjoy the consumption of a good that gives them satisfaction and enjoyment.
[For a random half of subjects, the previous two sentences are presented in reverse order]

For this question, assume that if the tax is introduced, Anne will reduce her consumption of sugary drinks so that her weight permanently drops to 145 pounds (from the previous 190 pounds). Given Anne's height, this is a "normal weight", according to the World Health Organization.

If all residents of Melbourne were exactly like Anne, would you support or oppose the introduction of the tax?
[I would strongly support the tax, I would weakly support the tax, I would weakly oppose the tax, I would strongly oppose the tax]

How do you think Anne is affected overall if the tax is introduced?
[Anne will be much better off with the tax than without, Anne will be a little better off with the tax than without, Anne will be equally well off with the tax as without, Anne will be a little worse off with the tax than without, Anne will be much worse off with the tax than without]

Science does not have a definitive answer about how much the introduction of a sugary drinks tax will change Melbourne residents' body weight. What do you believe are likely effects?

What do you believe is the chance that the tax will cause a typical moderately obese person to lose enough weight to be classified as merely overweight? I believe this is
[Extremely likely, Very likely, Somewhat likely, Somewhat unlikely, Very unlikely, Extremely unlikely]

A sugary drinks tax will not affect everyone equally, because some people are richer than others, some are heavier than others, some like sugary drinks more than others, and so on. On the left you see 10 tags labelled "Random resident". Please sort the labels into the bins to show, from amongst ten randomly selected Melbourne residents, how many people you believe would be affected in what way.
[Subjects drag and drop ten tags labelled "Random resident" into three bins labeled "Better off", "Neither better nor worse off", "Worse off"]

Effect on whole society
If residents like Anne have a healthier weight, this affects the well-being of these residents. It also affects the general public. The reason is that people who have a healthier weight are less likely to get ill and cause medical costs. Health insurance in Australia is partly publicly provided. Hence, the less a person relies on it, the lower the costs she causes to the public.

Suppose the city of Melbourne asked for your opinion on the proposed sugary drinks tax. Please indicate your opinion. I would
[Strongly support the proposed sugary drinks tax, Weakly support the proposed sugary drinks tax, Weakly oppose the proposed sugary drinks tax, Strongly oppose the proposed sugary drinks tax]

What determined your answer in the previous question? The well-being of those who change their sugary drinks consumption, or the effects on society at large (lower health costs)?
[Exclusively the change in well-being of those who change their sugary drinks consumption; Mostly the change in well-being of those who change their sugary drinks consumption; Both, but a little more the change in well-being of those who change their sugary drinks consumption; Both, but a little more the effects on society at large (lower health costs); Mostly the effects on society at large (lower health costs); Exclusively the effects on society at large (lower health costs)]

Questions for bonus payment About how tall is Anne?
[5 feet 0 inches, 5 feet 4 inches, 5 feet 8 inches, 6 feet 2 inches]

What effect would the tax have on the overall price of sugary drinks?
[It would increase one-and-a-half fold, It would increase two fold, It would increase two-and-a-half fold, It would increase three fold, It would increase three-and-a-half fold]

Policy 2 of 4: Alcohol taxes in London, UK

The next questions are all about Peter, an undergraduate student at Queen Mary University, London, UK. Like many undergraduate students, on one day of each weekend, Peter consumes five or more alcoholic drinks within two hours or less, a practice known as binge-drinking. The price of alcohol has a large influence on how much alcohol people consume. Suppose the United Kingdom considers increasing the alcohol tax, such that each unit of alcohol would become twice as expensive as it is today. The value added tax on all other goods would be lowered, so that the state would collect the same overall amount of taxes as without the change.
If the increased tax reduces alcohol consumption, this will have two effects. On the one hand, it will limit the unfavorable health effects of alcohol consumption. On the other hand, people will less frequently enjoy consuming something they like.
[For a random half of subjects, the previous two sentences are presented in reverse order]

For the following questions, assume that if the tax is introduced, Peter reduces his binge drinking from once a week to once a month.

If all residents of London were exactly like Peter, would you support or oppose the tax?
[I would strongly support the tax, I would weakly support the tax, I would weakly oppose the tax, I would strongly oppose the tax]

How do you think Peter is affected overall if the tax is introduced— *[Peter will be much better off with the tax than without, Peter will be a little better off with the tax than without, Peter will be equally well off with the tax as without, Peter will be a little worse off with the tax than without, Peter will be much worse off with the tax than without]*

Science does not have a definitive answer about how much the introduction of an alcohol tax will affect the frequency of binge drinking. What do you believe is the chance that the tax will reduce the frequency by which people like Peter binge-drink by at least half? I believe this is
[Extremely likely, Very likely, Somewhat likely, Somewhat unlikely, Very unlikely, Extremely unlikely]

An alcohol tax will not affect everyone equally, because some people are richer than others, some like to drink alcoholic drinks more than others, and so on. On the left you see 10 tags labelled "random college student". Please sort the labels into the bins to show how many college students you believe would be affected in what way.
[Subjects drag and drop ten tags labelled "Random resident" into three bins labeled "Better off", "Neither better nor worse off", "Worse off"]

Effect on whole society
If college students like Peter binge-drink less frequently, this affects the well-being of these students. It also affects the general public. The reason is that binge-drinking college students sometimes cause trouble to others, for instance, by vomiting in public places, through vandalism, or harassing bystanders.

Suppose the city of London asked for your opinion on the proposed alcohol tax. Please indicate your opinion. I would
[Strongly support the proposed alcohol tax, Weakly support the proposed alcohol tax, Weakly oppose the proposed alcohol tax, Strongly oppose the proposed alcohol tax]

What determined your answer in the previous question? The well-being of those who change their alcohol consumption, or the effects on society at large (less vomit, vandalism, etc.)?
[Exclusively the change in well-being of those who change their alcohol consumption; Mostly the change in well-being of those who change their alcohol consumption; Both, but a little more the change in well-being of those who change their alcohol consumption; Both, but a little more the effects on society at large (less vomit, vandalism, etc.); Mostly the effects on society at large (less vomit, vandalism, etc.); Exclusively the effects on society at large (less vomit, vandalism, etc.)]

Questions for bonus payment
How often does Peter binge-drink?
[Twice a week, Once a week, Twice a month, Once a month, Once every other month]

What effect would the tax have on the overall price of alcohol?
[It would increase one-and-a-half fold, It would increase two fold, It would increase two-and-a-half fold, It would increase three fold, It would increase three-and-a-half fold]

Policy 3 of 4: Restrictions on payday loans in Christchurch, New Zealand

The next questions are all about James, a 40 year old janitor at a school in Christchurch, New Zealand. James receives his paycheck monthly. In the middle of the month, James is running very tight on money. Suddenly, a need comes up that James considers very important. James considers taking out $500 from a payday loan store. If he does so, he will have to repay the $500 he has taken out, in addition to $90 in interest and fees, two weeks later. Currently, New Zealand does not have any restrictions on the interest and fees that payday lenders can charge their clients. New Zealand is considering introducing tight restrictions on payday lending. Those restrictions would consist of upper limits on the interest rates and fees that payday lenders can charge.
Such regulation would have two effects. First, people can get trapped in debt cycles in which they take up loans to pay back previous loans, at ever increasing interest and fee payments. Some experts argue that restrictions on payday lending make debt cycles less likely. Second, fewer people will be able to get a payday loan, because some payday loan shops may go out of business. Hence, some people who would genuinely benefit from a payday loan may no longer be able to obtain them.
[For a random half of subjects, the previous two arguments are presented in reverse order]

Suppose regulation is introduced such that James will be unable to get the payday loan of $500 for a cost of $90 in interest and fees two weeks later. Suppose that James will also be unable to borrow that money from anywhere else.

If all residents of New Zealand were exactly like James, would you support or oppose the introduction of the restrictions?
[I would strongly support the restrictions (so loans become unavailable), I would weakly support the restrictions (so loans become unavailable), I would weakly oppose the restrictions (so loans remain available), I would strongly oppose the restrictions (so loans remain available)]

How do you think James is affected if the regulation is introduced?
[James will be much better off if he cannot get the loan, James will be a little better off if he cannot get the loan, James will be just as well off if he can get the loan as if he cannot get it, James will be a little better off if he can get the loan, James will be a much better off if he can get the loan]

Science does not have a definitive answer about exactly how much the introduction of payday loan regulation limits individuals' ability to obtain high-cost, short-term credit. The reason is that even people they cannot get a payday loan, they may find alternatives, such as pawn shops, bank overdrafts, or asking friends and family. All of these options may also cost much in interest and fees or damaged social relationships, and bear a risk of trapping people in debt cycles.

What do you think are likely effects? I believe that if James can no longer get a payday loan, the chance that he will get a loan from some other place at similar costs with similar risks is
[Extremely high (>90%), Very high (75% - 90%), Somewhat high (50% - 75%), Somewhat low (25%-50%), Very low (10%-20%), Extremely low (<

10%)]

Restrictions on payday lending will not affect everyone equally, because people differ in their reasons for taking out these loans. On the left you see 10 tags labelled "Random resident". Please sort the labels into the bins to show how many people like James (who would otherwise get a payday loan) you believe would be affected in what way if payday loans were no longer available.
[Subjects drag and drop ten tags labelled "Random resident" into three bins labeled "Better off", "Neither better nor worse off", "Worse off"]

Effect on whole society
If workers like James can no longer take out payday loans, this affects these workers. It also affects the general public, for instance by changing the probability that people go bankrupt or become homeless. Suppose the country of New Zealand asked for your opinion on the proposed regulation to limit the availability of payday loans. Please indicate your opinion. I would
[Strongly support the proposed regulation, Weakly support the proposed regulation, Weakly oppose the proposed regulation, Strongly oppose the proposed regulation]

What determined your answer in the previous question? The well-being of those who change their borrowing behavior, or the effects on society at large?
[Exclusively the change in well-being of those who change their borrowing behavior; Mostly the change in well-being of those who change their borrowing behavior; Both, but a little more the change in well-being of those who change their borrowing behavior; Both, but a little more the effects on society at large; Mostly the effects on society at large; Exclusively the effects on society at large]

Questions for bonus payment

If James took out a payday loan of $500 today, how much would he have to repay in total (including interest and fees) in two weeks?
[$500 - $600 in steps of $10]

What line of work is James in?
[Factory worker, Farmer, Janitor, Truck driver, Server at a restaurant, Musician]

Policy 4 of 4: Minimal retirement savings requirements in Dublin, Ireland
The next questions are all about Frank, a 30 years old resident of Dublin, Ireland. Frank's income is EUR2900 per month, which is about USD3300 per month. Currently, Frank does not save anything for retirement. Research finds that in countries in which there is no requirement that people save for retirement, such as the US, only about 1 in 5 working age adults have a retirement savings plan, and that many will start retirement with savings far lower than needed to maintain their standard of living when transitioning into retirement. Ireland currently does not have a minimum retirement savings mandate. For this question, suppose that Ireland considers introducing a law that every working-age adult must save at least 10% of their income for their retirement. Financial advisors recommend that everyone save between 10% and 30% of their income for retirement, depending on their age. If the policy is introduced, Frank would be forced to pay at least $330 per month into a retirement savings account. Upon retirement, he would receive a constant monthly pension for the rest of his life. With the policy could expect to experience only a moderate drop in his standard of living when transitioning into retirement.
Some therefore argue that people should be forced to save some minimal amount for retirement to ensure nobody experiences a large drop in their standard of living upon retirement. Others argue that people should be free to decide how and when to spend or save their money.
[For a random half of subjects, the previous two arguments are presented in reverse order]

Suppose the law is passed that requires Frank to save at least 10% of his income for retirement ($330 of the $3300 he earns each month). If all residents of Ireland were exactly like Frank, would you support or oppose the introduction of the minimal retirement savings law?
[I would strongly support the minimal retirement savings law, I would weakly support the minimal retirement savings law, I would weakly oppose the minimal retirement savings law, I would strongly oppose the minimal retirement savings law]

How do you think Frank is affected if the regulation is introduced?
[Overall, Frank will be much better off if he is forced to save; Overall, Frank will be a little better off if he is forced to save; Overall, Frank will be just as well off if he is forced to save as if he is not; Overall, Frank will be a little worse off if he is forced to save; Overall, Frank will be much worse off if he is forced to save]

Minimal retirement savings laws will not affect everyone equally, because people differ in their financial backgrounds. Some are in stable jobs, while others may be in a period of their lives in which they earn unusually little or unusually much. On the left you see 10 tags labelled "Random resident". Please sort the labels into the bins to show how many people in Ireland you believe would be affected in what way if a minimal compulsory retirement savings rate of 10% were introduced.
[Subjects drag and drop ten tags labelled "Random resident" into three bins labeled "Better off", "Neither better nor worse off", "Worse off"]

Effect on whole society

If workers like Frank save differently for retirement, this affects the well-being of these workers. It also affects the general public, for instance by changing the probability that people go bankrupt or become homeless. Suppose the Republic of Ireland asked for your opinion on the proposed minimal retirement savings law. Please indicate your opinion. I would
[Strongly support the law that forces workers to save 10% of their income for retirement, Weakly support the law that forces workers to save 10% of their income for retirement, Weakly oppose the law that forces workers to save 10% of their income for retirement, Strongly oppose the law that forces workers to save 10% of their income for retirement]

What determined your answer in the previous question? The well-being of those who change their saving behavior, or the effects on society at large (reliance on social security)?
[Exclusively the change in well-being of those who change their saving behavior; Mostly the change in well-being of those who change their saving behavior; Both, but a little more the change in well-being of those who change their saving behavior; Both, but a little more the effects on society at large; Mostly the effects on society at large; Exclusively the effects on society at large]

Questions for bonus payment

What is Frank's approximate monthly income (in USD)
[$1000 - $10,000 in steps of $1000]

If the minimal retirement savings law is introduced, what percentage of his earnings will each resident have to put away, at least, each month?
[1%, 2.5%, 5%, 7.5%, 10%, 12.5%, 15%, 17.5%, 20%, 25%, 30%]

Questions about yourself

As the last part of this survey, we would like to ask 22 questions about yourself. Please answer truthfully.

What is your age?
[18 - 90 in steps of 1 year]

What is your gender?
[male, female, other (e.g. non-binary)]

How do you feel about your own alcohol consumption?
[I am perfectly comfortable with my alcohol consumption; I am fairly comfortable with my alcohol consumption; I feel neutral about my alcohol consumption; I am fairly uncomfortable with my alcohol consumption; I am totally uncomfortable with my alcohol consumption]

How do you feel about your use of short-term credit (such as payday loans)?
[I am perfectly comfortable with my use of short-term credit; I am fairly comfortable with my use of short-term credit; I feel neutral about my use of short-term credit; I am fairly uncomfortable with my use of short-term credit; I am totally uncomfortable with my use of short-term credit]

How do you feel about your retirement savings choices?
[I am perfectly comfortable with my retirement savings choices; I am fairly comfortable with my retirement savings choices; I feel neutral about my retirement savings choices; I am fairly uncomfortable with my retirement savings choices; I am totally uncomfortable with my retirement savings choices]

How do you feel about your body weight?
[I am perfectly comfortable with my body weight; I am fairly comfortable with my body weight; I feel neutral about my body weight; I am fairly uncomfortable with my body weight; I am totally uncomfortable with my body weight]?

How many alcoholic beverages did you consume on average per week, calculated over the last 12 months? (1 alcoholic drink = 0.2 liter beer, 0.1 liter wine, 1 shot of schnapps or liqueur)
[10 or more per week, 5-10 per week, 3-5 per week, 1-3 per week, less than 1 per week but some, none]

How often have you had 4 or more alcoholic beverages within a 2-hour period over the last 12 months?
[For males, "4 or more" is replaced by "5 or more"]
[7 days per week, 5-6 days per week, 3-4 days per week, 2 days per week, 1 day per week, 2-3 days per month, 3-11 days over the last year, 1-2 days over the last year, never]

What is your body height?
Feet *[Subject can enter any integer]*
Inches *[Subject can enter any integer]*

What is your body weight (in pounds)? *[Subject can enter any integer]*

How would you describe yourself?
[Underweight, Healthy weight, Overweight, Moderately obese, Severely obese, prefer not to answer]

Please indicate your current household income in U.S. dollars
[Under $10,000, $10,000 - $19,999, $20,000 - $29,999, $30,000 - $39,999, $40,000 - $49,999, $50,000 - $74,999, $75,000 - $99,999, $100,000 - $150,000, Over $150,000, prefer not to answer]

What is your credit card debt (across all credit cards you have)?
[$0-$500, $500-$1000, $1000-$2500, $2500-$5000, $5000-$7500, $7500-$10,000, $10,000-$20,000, $20,000-$50,000, $50,000 or more, prefer not to answer]

Have you ever taken a payday loan?
[No, never; Yes, once; Yes, a couple of times; Yes, often; prefer not to answer]

Have you ever been in a debt cycle (getting into debt in order to repay other debt)?
[No, never; Yes, once; Yes, a couple of times; Yes, often; prefer not to answer]?

What are your total retirement savings?
[$0, $0 - $10,000, $10,000-$25,000, $25,000-$50,000, $50,000-$100,000, $100,000-$200,000, $200,000-$300,000, $300,000-$400,000, $400,000-$500,000, $500,000-$750,000, $750,000-$1,000,000, more than $1,000,000, I do not know, prefer not to say]

Approximately what percentage of your income do you currently save for retirement each month?
[0%, 0%-2.5%, 2.5%-5%, 5%-7.5%, 7.5%-10%, 10%-15%, 15%-20%, 20%-30%, more than 30%]

Do you currently work for a company that offers a retirement savings plan (401k)?
[Yes, No]

Where do you stand politically?
[Conservative; Leaning conservative; Centrist; Leaning liberal; Liberal]

Please indicate the highest level of education you completed.
[Elementary School, Middle School, High School or equivalent, Vocational/Technical School (2 year), Some College, College Graduate (4 year), Master's Degree (MS), Doctoral Degree (PhD), Professional Degree (MD, JD, etc.)]

Which of the following best describes the area you live in?
[Urban, Suburban, Rural]

Please choose the option that best describes your situation
[I am unemployed, I am employed part-time, I am employed full-time]