European Economic Review 176 (2025) 105022

Contents lists available at ScienceDirect

European Economic Review

journal homepage: www.elsevier.com/locate/eer

Freedom counts: Cross-country empirical evidence

João V. Ferreira [a], Nobuyuki Hanaki [b,c], Fabrice Le Lec [d], Erik Schokkaert [e], Benoît Tarroux [f,*]

[a] University of Southampton, Building 58, Highfield Campus, SO17 1BJ, Southampton, UK
[b] Institute of Social and Economic Research, Osaka University, 6-1 Mihogaoka, Ibaraki, Osaka, 567-0047, Japan
[c] University of Limassol, Cyprus
[d] Univ. Lille, CNRS, IESEG School of Management, UMR 9221 LEM Lille Économie Management, 42 rue Paul Duez, F-59000, Lille, France
[e] Department of Economics, KU Leuven, Naamsestraat 69, B-3000 Leuven, Belgium
[f] Université Lumière Lyon 2, CNRS, Université Jean Monnet Saint-Etienne, emlyon business school, GATE, 69007, Lyon, France

ARTICLE INFO

Dataset link: Replication package for "Freedom counts: Cross-country Empirical Evidence" (Original data)

JEL classification:
C83
D63
I31

Keywords:
Freedom of choice
Welfare
Intrinsic value
Opportunity set
Cross-cultural survey

ABSTRACT

We investigate whether people attach intrinsic value to freedom and which theoretical rules they implicitly employ to evaluate opportunity sets. To do this, we run a new survey-based study with 4902 participants across 10 different countries. Participants face various comparisons of opportunity sets in a policy-relevant context in a "spectator" position. Our main result is that an overwhelming majority of spectators *reveal* that they attach intrinsic value to freedom. We also find that a large majority of participants use size-based rules to rank sets in terms of freedom, while there is considerable heterogeneity in the theoretical rules they implicitly employ to rank sets in terms of welfare. These results are strikingly robust across countries.

1. Introduction

How to rank *opportunity sets* — the set of options available to individuals — is an important normative and empirical question, with implications for organizations, markets, and public policy. In economics, it is standard practice to rank opportunity sets based on the "best" alternatives available on those sets, where "best" is defined on the basis of individual preferences. According to this approach (called *indirect utility*, or IU for short), freedom of choice (hereafter FoC) has only *instrumental value*: additional alternatives in an opportunity set are valuable only if they lead to the choice of a better alternative in the individual's preference ranking. Although this approach is standard and parsimonious, two important empirical questions remain unanswered.

First, do people only attach instrumental value to freedom, or do they also attach *intrinsic value* to it? The literature on FoC — as pioneered by Amartya Sen (1985, 1988) and developed by many others — postulated that additional alternatives can be valuable even if they do *not* lead to the choice of a better alternative.[1] From a policy perspective, understanding individual attitudes towards

* Correspondence to: Univ. Lumière Lyon 2, GATE, 35, Rue Raulin, 69007 Lyon, France
E-mail addresses: j.ferreira@soton.ac.uk (J.V. Ferreira), nobuyuki.hanaki@iser.osaka-u.ac.jp (N. Hanaki), fabrice.lelec@univ-lille.fr (F. Le Lec), erik.schokkaert@kuleuven.be (E. Schokkaert), benoit.tarroux@univ-lyon2.fr (B. Tarroux).

https://doi.org/10.1016/j.euroecorev.2025.105022
Received 9 August 2024; Received in revised form 20 March 2025; Accepted 21 March 2025
Available online 23 April 2025

FoC is important to evaluate the legitimacy and potential behavioral effects of a large range of public policies, such as paternalistic and soft paternalistic interventions. As such, there is a growing literature on individual attitudes towards paternalistic interventions (e.g., Ambuehl et al. 2021, Dreyer and Mahler 2022, Bartling et al. 2023) and the trade-off between satisfying preferences and having freedom (e.g., Alsan et al. 2023). However, to the best of our knowledge, there is little if any empirical evidence on whether people attach intrinsic value to freedom.

A second related and important question, for which there is also very limited evidence, is to know how people compare different opportunity sets in terms of freedom and welfare.[2] Which theoretical ranking rules, like IU, do people (implicitly) employ when ranking opportunity sets? This is again an empirical question. It is distinct from the normative question of how we *should* rank opportunity sets (see Barberà et al. 2004, Foster 2011 for reviews). However, following the tradition of *positive welfare economics*, we argue that it can inform the latter question (see, e.g., Gaertner and Schokkaert 2012, Ambuehl and Bernheim 2021). For example, understanding how people rank opportunity sets can bring valuable insights into the political feasibility and democratic desirability of different theoretical ranking rules.

In this paper, we use a novel survey-based research design to (i) understand whether people attach intrinsic value to FoC in a policy-relevant context, and (ii) investigate which theoretical rules people implicitly employ when ranking opportunity sets in terms of FoC and overall welfare. We do this for a total of 4902 participants across 10 different countries.[3] Participants face 15 comparisons of opportunity sets in a health-related context in a "spectator" position (i.e., they compare menus of hospitals for two identical individuals in a similar position to a policymaker).[4] For each *set comparison*, we ask them to evaluate which set provides more FoC (our Freedom question), what is the "best" alternative present in the sets (our IU question), and which set is best overall (our Welfare question). To investigate whether spectators attach intrinsic value to FoC, we look at set comparisons where subjects respond that there is a "conflict" between FoC and the "best" alternative (i.e., when a set A has more FoC and the best alternative is in a set B or in both sets). Then, if a participant states that set A is best overall, we say that this participant *reveals* giving intrinsic value to freedom. To investigate which rules people use, we draw on the theoretical literature and identify a reasonably large number of plausible theoretical rules that can be used to rank sets in terms of FoC and welfare. Each of these rules implies a distinctive *theoretical response pattern* for the 15 set comparisons. We then compare this theoretical response pattern to the *actual response pattern* of participants using a Bayes classification procedure, in order to assign participants to the theoretical rule that best matches their responses in the Freedom question and the (potentially different) rule that best matches their responses in the Welfare question.

We find that an overwhelming majority of spectators (> 80%) *reveal* that they attach intrinsic value to freedom at least once. In addition, participants attach intrinsic value to freedom in 48% of the relevant cases. Regarding our second question, we find, first, that the overwhelming majority of subjects respond *as if* they rely on cardinality-based rules (rules that "count" the number of alternatives) to rank sets in terms of FoC. Second, when ranking sets in terms of overall welfare, we find that quality-based rules (such as IU or other rules that take into consideration the "quality" of alternatives) dominate, but there is considerable heterogeneity in the rules that subjects implicitly employ. In terms of country comparisons, we find that our results are strikingly similar across countries with very different social, cultural, and institutional backgrounds. This suggests that people's attitudes toward freedom, as well as their way of comparing opportunity sets, are robust cross-cultural patterns. Finally, we show that the fit of our classification results is overall good, that we have not omitted empirically important rules, and that our main results for the classification and intrinsic value of freedom are not driven by random answers, mistakes or intuitive responses, salience effects, preferences for flexibility, or other potential confounding factors.

Our analysis is important from both an academic and a policy perspective. From an academic point of view, our analysis contributes to the growing literature on individual attitudes towards paternalistic interventions (e.g., Arad and Rubinstein 2018, Ackfeld and Ockenfels 2021, Ambuehl et al. 2021, Dreyer and Mahler 2022, Bartling et al. 2023) and the trade-off between satisfying preferences and having freedom (e.g., Alsan et al. 2023). We show that the intrinsic value of freedom should be taken into account when evaluating these policies and trade-offs. We are also the first to provide empirical evidence on the theoretical rules that people implicitly employ when evaluating menus of options in terms of freedom and welfare. This provides important empirical facts to determine the plausibility, democratic desirability, and political feasibility of the ranking rules proposed by the theoretical/normative literature cited above. Note that such empirical evidence can have important implications for how policymakers design their choice environments, a point we elaborate on below. Our results for 10 countries also offer crucial insights into the evaluation of opportunity sets for citizens from a large variety of social, cultural, and institutional backgrounds. In particular, our results uncover attitudes toward freedom that appear robust across countries. Methodologically, our analysis of the intrinsic value of freedom offers a clear empirical test that can inform future research.

---

[1] See Sugden (1998), Barberà et al. (2004), Baujard (2007), Gravel (2008), Dowding and van Hees (2009), and Foster (2011) for reviews. Why would individuals value freedom beyond its instrumental benefits? Possible underlying reasons are that FoC, independently of the alternatives that are chosen, is an element of a person's well-being (e.g., Sen 1988), is important to develop mental faculties such as judgment and self-control (e.g., Mill 1859, p. 117), or allows people to lead autonomous and therefore meaningful lives (e.g., Nozick 1974, pp. 48–51).

[2] "Welfare" is here understood as an all-things-considered evaluation of sets (see Hausman 2012). In other words, everything that contributes to the value of an opportunity set is taken into account for welfare.

[3] Our list of countries includes Brazil, China, Colombia, France, Japan, Netherlands, Portugal, Turkey, the United Kingdom, and the US.

[4] There is a large and growing literature in economics that uses a spectator position in vignette and experimental studies (e.g., Konow 2009, Almås et al. 2020, Müller and Renes 2021). We further justify the relevance of this setting below.

In addition, our paper contributes to the growing empirical literature on choice over menus. Following seminal theoretical contributions to the study of preferences over menus (e.g., Kreps 1979, Gul and Pesendorfer 2001), several lab experimental studies have sought to elicit these preferences (e.g., Toussaert 2018, Le Lec and Tarroux 2020, Arlegi et al. 2022). Interestingly, Le Lec and Tarroux's (2020) lab experiment provides evidence that subjects value *less* a menu than its preferred option in a consumption environment. At first sight, this result seems to contradict our finding about the intrinsic value of freedom. Note, however, that the two studies differ in important aspects. In particular, in Le Lec and Tarroux (2020) constraints on freedom are endogenously imposed by the subjects themselves. In addition, the trade-off between the value of freedom and psychological costs associated with greater freedom are more salient in their setting. For that reason, their result can be rationalized by subjects' fear of making mistakes in the presence of a larger menu. In our study, constraints on freedom are exogenously imposed and we elicit people's preferences in a spectator position. We find evidence that some spectators judge smaller menus to be better when others can commit a (very) costly mistake, but the dominant pattern is for larger menus to be considered better even if they have the same (or worst) top alternative than smaller menus.[5] Therefore, we contribute to this growing literature by being the first to measure whether attitudes toward freedom of choice influence spectator preferences over menus in a policy-relevant (health-related) context. Overall, our results show that the intrinsic value of freedom of choice substantially influences spectators' preferences over menus and that IU will disregard relevant welfare-enhancing features of opportunity sets (see Benjamin et al. 2014 for related evidence).

From a policy perspective, the attitudes of the population on social issues can be an important input into the process of democratic public decision-making. If (as we find) a large share of the population thinks that freedom has intrinsic value in policy-relevant contexts, this implies that people's attitudes toward menus can complement information about preferences over alternatives as inputs to public policy. If policymakers wish to respect people's attitudes toward menus, our results can also inform them on how to design various choice environments. For instance, our results strongly suggest that for a substantial part of the population, it is important to enlarge small opportunity sets even if this does not lead to a change in behavior. This is an important finding since many policy decisions are between providing a single option or a small number of options, as is the case in many countries for how many schools people can choose for their children, how many hospitals they can choose to receive care, or how many pension schemes they can choose in their companies. Our study is set in the context of hospital choice and provides direct evidence relevant to ongoing debates about people's freedom of choice in this setting (see, e.g., UK Government 2023). As discussed later, we designed our study to enhance the generalizability of our findings to similar contexts where the number of available options is small.

The remainder of the paper is organized as follows. In the next section, we formally define the revelation mechanism for the intrinsic value of freedom of choice and we present the theoretical ranking rules that we test in our data. In Section 3, we present the research design of the study. Section 4 is devoted to the empirical analysis and results, and Section 5 concludes.

2. Conceptual background

To illustrate the problem, consider the following simple example:

- **Opportunity set A:** (apple, orange)
- **Opportunity set B:** (apple)

where "apple" and "orange" are possible alternatives for consumption and opportunity sets A and B are menus from which those fruits can be chosen. In this section, we formally define what it means for an individual to reveal that he/she attaches intrinsic value to FoC in such situations. We also summarize the main rules that have been proposed in the theoretical literature for ranking different opportunity sets in terms of freedom and welfare.

Before proceeding, we introduce a bit of notation. Let X be a finite set of alternatives, denoted below by x, y, and z, and let A and B denote non-empty subsets of X (i.e., opportunity sets). Let $\succsim^F$ be a transitive quasi-ordering over opportunity sets such that $A \succsim^F B$ means that "A provides at least as much freedom of choice as B", with $\sim^F$ and $\succ^F$ being the symmetric and asymmetric components of $\succsim^F$. We denote by $\succsim^W$ a transitive quasi-ordering over opportunity sets such that $A \succsim^W B$ means that "A provides at least as much welfare as B", with $\sim^W$ and $\succ^W$ being the symmetric and asymmetric components of $\succsim^W$. Our interpretation of this welfare relation is an *all-things-considered* comparison of sets. Let $\succsim_i^W$ and $\succsim_i^F$ correspond to $\succsim^W$ and $\succsim^F$ as judged by individual i. Finally, let R_i denote individual i's transitive and complete preference ordering over feasible alternatives (with P_i and I_i its asymmetric and symmetric components) and $\max(A)$ denote individual i's set of preferred element(s) of A such that $\max(A) = \{x : xR_i y \ \forall y \in A\}$. With a slight abuse of notation, we write $\max(A) R_i \max(B)$ to denote that individual i weakly prefers the best element(s) in A to those in B, with $\max(A) P_i \max(B)$ for strict preference and $\max(A) I_i \max(B)$ for indifference.

2.1. The intrinsic value of freedom of choice

How can an individual *reveal* that he/she attaches intrinsic value to freedom of choice (hereafter also IvFoC) through their choices? Formally:

[5] In line with this result, previous lab experiments using a principal–agent situation have shown that subjects are willing to pay for the intrinsic value of choosing for themselves when interacting with others or with bots can reduce their decision rights (e.g., Bartling et al. 2014, Ferreira et al. 2020).

Definition 1 (*IvFoC Revelation Principle*). For two sets A and B such that $A \succ_i^F B$ and $\max(B) R_i \max(A)$, individual i reveals attaching IvFoC if $A \succ_i^W B$.

In other words, we say that i reveals that he/she attaches IvFoC whenever faced with sets A and B such that i judges that (i) A provides more FoC than B and (ii) the best alternative in B is at least as good as the best alternative in A, individual i considers that A provides more overall welfare than B. For example, take the opportunity sets $A = \{x, y, z\}$ and $B = \{x, y\}$. Assume that i considers that set A provides more FoC than set B. If i deems that, all things considered, set A provides more welfare than set B even though the best alternative in both sets is x (i.e., $\max(B) I_i \max(A)$), then i responds *as if* he/she attaches intrinsic value to having more FoC. Now consider $A = \{x, y, z\}$ and $B = \{w\}$. A stronger case of IvFoC is if i considers that A provides more welfare and FoC than set B, even though w is better than x, y, and z (i.e., $\max(B) P_i \max(A)$). Are there individuals who give so much intrinsic value to freedom that they prefer opportunity sets with more freedom but with a "worse" top alternative?

2.2. Ranking opportunity sets

The theoretical literature on ranking opportunity sets is vast. Nonetheless, it is possible to organize the existing theoretical ranking rules in different *families*. In this paper, we focus on three main families: the *cardinality family*, focusing on the "size" of the menu (Section 2.2.1), the *indirect utility family*, ranking the sets based on the perceived "quality" of their best alternatives (Section 2.2.2), and the *potential preferences family*, using multiple "reasonable" preference orderings for the evaluation of sets (Section 2.2.3). In Section 2.2.4, we discuss some other theoretical rules that we test in the data.

We cover rules that have been proposed to rank sets in terms of FoC, welfare, or both. Note, however, that all rules can *a priori* be used to rank sets in terms of FoC or welfare. In our empirical analysis, we will test the significance of all the rules we cover in both dimensions. Overall, we test a reasonably large set of plausible ranking rules (Appendix A provides a summary of the 23 rules we consider and their formal definitions; all appendices available online).[6]

2.2.1. Cardinality family

The first family of rules focuses on the size of the opportunity sets. A seminal ranking rule, the *cardinality rule*, ranks opportunity sets based on the number of alternatives of each set (Pattanaik and Xu, 1990). This rule focuses on the "quantity of action" available to a person (Carter, 1999) and may be "a natural way of measuring freedom in the absence of information about the agent's preferences" (Foster 2011, p. 20).[7]

The cardinality rule can be weakened to the so-called *weak cardinality rule* (see Puppe 1996 for a related rule). This rule still compares sets based on their size, but states that adding an alternative to a set *will never decrease* (as opposed to *will always increase*) freedom/welfare. As part of this family, we also consider a *diversity rule* that counts the number of "non-similar options" in a set. It says that adding a non-similar option to a set always increases freedom/welfare, while adding an option to a set that is similar to another option available in that set does not increase freedom/welfare.[8]

2.2.2. Indirect utility family

Cardinality-based rules are often criticized for not taking into account the "quality" of the alternatives or the elements present in the sets (e.g., Sen 1990, Sugden 1998). For example, according to cardinality-based rules, a singleton menu A with the option "rotten apple" and a singleton menu B with the option "good apple" should be equally ranked.

An alternative is then to use rules that take the quality of alternatives into consideration. This is particularly relevant if one is ranking opportunity sets in terms of welfare, but quality-based ranking rules have been proposed as freedom rules as well (e.g., Foster 2011).

Among these, the most prominent is the standard economics approach. Known as the *indirect utility rule*, it ranks sets based on their "best" element, where "best" is defined according to the individual's *actual preference* over alternatives. According to this rule, the menu with the best element for individual i offers greater freedom/welfare.

We also consider quality-based rules that abstract from preferences and define the quality of the elements on a cardinal scale. These rules describe alternatives by their quality in n attributes. The quality of x can then be written as $\sum_j^n w_j x_j$, where x_j is the quality of x in attribute j and w_j is a quality weight that can be varied as a sort of sensitivity analysis. In our empirical setting, alternatives are described by two attributes. Therefore, plausible variations of IU put more or less weight on one of the two attributes. In accordance, we test three *MaxMax rules* that weight the best alternative in one of the two attributes or weight the best alternative giving the same weight to both attributes [i.e., with weights $(w_1, w_2) = (0, 1)$ or $(w_1, w_2) = (1, 0)$ and $(w_1, w_2) = (0.5, 0.5)$ respectively].

[6] We demonstrate that these rules are exhaustive within our context. In Appendix D.5, we conduct a cluster analysis to identify patterns in subjects' responses, confirming that our main analysis does not overlook any significant rules. It is important to note that we do not consider *negative freedom* (e.g., Hayek 1960, Van Hees 1998) and other forms of interpersonal freedom (e.g., Sher 2018) that are unimportant (both theoretically and empirically) in our opportunity set context where interpersonal constraints are fixed, but that may affect the evaluation of social states in other settings.

[7] In their seminal paper, Pattanaik and Xu (1990) show how this ranking rule is equivalent to a set of simple axioms. For the sake of concision, we focus on the rules without introducing their equivalent axioms. Doing otherwise would greatly enlarge the conceptual background and divert attention from our main analysis.

[8] There are several competing proposals to measure diversity/similarity (e.g., Pattanaik and Xu 2000, Nehring and Puppe 2002, Bervoets and Gravel 2007). We sidestep this issue in our empirical setting by having options that are clearly similar and options that are clearly dissimilar without reference to preferences.

We restrict our analysis to this set of weights because in our empirical setting there are few ranking differences between these three weights *and* all other potential weights.

We also consider lexicographic rules that combine IU and cardinality, as proposed by Bossert et al. (1994). For example, the *Lex IU-Cardinality rule* first compares sets according to IU, and only in case of indifference based on IU, it then compares sets according to cardinality. This rule relies on "quality" to make judgments, but when two sets have the same quality level, it uses "quantity" to break the tie. The *Lex Cardinality-IU rule* works the other way around. In our analysis, we attach the *Lex IU-Cardinality rule* to the indirect utility family and the *Lex Cardinality-IU rule* to the cardinality family.

2.2.3. Potential preferences family

The third family we consider appeals to *potential preferences* to take the quality of alternatives into account, where these potential preferences are defined as "the range of preferences that the individual might have had in the relevant circumstances" (Sugden 1998, p. 323). For example, potential preferences can be all preferences that according to the observer can be "reasonably" held by individuals in the context of interest. For example, in most contexts it seems reasonable to prefer to "eat a good apple" than to "eat a good orange", or *vice-versa*; however, it does not seem ever reasonable to prefer to "eat a rotten apple" than to "eat a good apple". In that case, we say that "eat a good apple" and "eat a good orange" are *eligible* options, while "eat a rotten apple" is an *ineligible* option.

A prominent ranking rule in this family is the *range of opportunity rule*, according to which opportunity set A offers more freedom/welfare than B if A caters to more potential preferences than B (Pattanaik and Xu 1998, Sugden 1998). More precisely, it says that $A \gtrsim^{F} B$ (or $A \gtrsim^{W} B$) if and only if the number of *eligible* options in A that are at least as good as all the elements in B according to at least one potential preference is greater than the number of equivalent options in B.

Another ranking rule belonging to the potential preferences family is the *effective freedom rule* put forward by Foster (2011) (see Arrow 1995, Sen 2002 for related rules). This rule judges one set to have greater or equal freedom/welfare than another set if all potential preferences agree this is so. This ranking rule is "incomplete", as it does not compare sets for which not all potential preferences agree. Finally, we also consider lexicographic rules that combine the range of opportunity and cardinality rules. In our analysis, the *Lex RoO-Cardinality rule* belongs to the potential preferences family and the *Lex Cardinality-RoO rule* belongs to the cardinality family. We do not consider lexicographic rules that combine IU, range of opportunity, and effective freedom, because these are similar quality-based rules.

2.2.4. Other rules

The three families described in the previous sections contain most of the best-known rules to rank opportunity sets. We now describe some other prominent rules that we test in our data.

First, we look at a *significant choice rule*, according to which opportunity set A offers more freedom/welfare than B if A has more eligible "non-similar" options than B (Sugden 1998, Pattanaik and Xu 2000). This rule takes into account both quality (eligibility) and diversity (non-similarity) as defined above. Second, we consider rules that are "opposite" to some of the previous rules. The first of these, the *choice aversion rule*, exhibits a preference for smaller sets when the best alternative is the same. This can be seen as a lexicographic rule that combines IU with "anti-cardinality". The second of these is the *MaxMin rule*, which ranks set A better than B if A has a better "worst" alternative than B with quality weights $(w_1, w_2) = (0, 1)$, $(w_1, w_2) = (1, 0)$, or $(w_1, w_2) = (0.5, 0.5)$. The third is a *MaxAverage rule*, which ranks sets according to the average quality of the attributes of all options in those sets with quality weights $(w_1, w_2) = (0, 1)$, $(w_1, w_2) = (1, 0)$, or $(w_1, w_2) = (0.5, 0.5)$. The fourth is the *intersection rule* — proposed by Bossert et al. (1994) — that only ranks sets for which IU and cardinality agree: a set offers more freedom/welfare than another set if it offers more freedom/welfare according to both rules. In contrast to the lexicographic rules that seek to find a "balance" when the two rules disagree, the intersection rule does not offer a ranking of the sets in such situations; instead, it is incomplete. Finally, we test a *trivial rule* that states that all opportunity sets offer the same degree of freedom/welfare. This rule is reviewed in Foster (2011), and it is used here mainly as a robustness check.

3. Research design

In this section, we present our survey-based research design. Survey-based studies have become increasingly popular in economics as they provide insights into attitudes, perceptions, and beliefs that are difficult to reveal through choice-based methods.[9] In our study, this approach enables us to examine attitudes toward freedom and welfare in a policy-relevant context that would be unfeasible to replicate in the lab. Moreover, findings from incentivized experiments — typically derived from consumption choices — may not generalize to contexts such as hospital choice, which is our setting. Therefore, we view our study as complementary to incentivized (lab) experiments, which could explore, for instance, the behavioral underpinnings of the intrinsic value of freedom. Additionally, it is worth noting that typical concerns with non-incentivized data are either not salient in our context (e.g., social desirability bias) or are addressed in our analysis (e.g., measurement error). Finally, this approach allowed us to gather data from a large, diverse sample of the general population across multiple countries in a cost-effective manner.

[9] See Gaertner and Schokkaert 2012, Stantcheva 2023, Haaland et al. 2023 for reviews. Recent studies have shown that survey measures can provide robust results consistent with incentivized experimental tools in various domains including risk, time, competition, and distributional preferences (Fallucchi et al., 2020; Bauer et al., 2020; Hufe and Weishaar, 2024).

3.1. Context

Participants are first presented with a short *vignette*, i.e., a brief description of a hypothetical scenario (see Appendix G for the full instructions). In the vignette, participants are told that two individuals, called Mr. Green and Mr. Yellow, who *are identical in all respects*, have to undergo a surgical procedure that is of minimal risk to their overall health.[10] However, this procedure requires them to spend four days recovering in a hospital and they have to choose a hospital for this surgery and for the recovery time. Importantly, the two individuals have different opportunity sets. Participants are told that their opportunity sets (the hospitals they can choose from) depend on their health insurance plans.[11]

Participants face 15 *set comparisons*. In a typical set comparison, the participants have to compare two sets of hospitals, in which hospitals differ in terms of *staff quality* ("service and assistance quality, nursing quality, friendliness of staff, etc., excluding doctors") and *comfort quality* ("bed quality, food quality, amenities, etc".).[12] Each attribute is rated from 1% for the lowest quality to 100% for the highest quality (we tell participants that the ratings are from "a trustworthy non-government agency that rates hospitals in their country"). The following set comparison is an example from the survey:

- Mr. Green has the following hospitals available in his insurance plan:
  - Hospital A (staff 80%, comfort 75%)
- Mr. Yellow has the following hospital available in his insurance plan:
  - Hospital A (staff 80%, comfort 75%)
  - Hospital B (staff 71%, comfort 89%)

We describe hospitals using two attributes for two main reasons. First, it allows us to test whether participants employ rules from the potential preferences family when ranking opportunity sets, a prominent set of rules in the theoretical literature. This is so because these rules only differ from IU in set comparisons where no dominant alternative exists, which is made possible by describing alternatives with two attributes. Second, using two attributes enables us to test theoretical ranking rules that take the diversity (or similarity) of alternatives into account.

3.2. Main questions

For each set comparison, participants responded to the following three questions (displayed on the same screen without the labels in bold):

- **Q1 (Freedom question):** *Which individual do you think has more freedom of choice?* [Answer options: Mr. Green/Mr. Yellow/The same]
- **Q2 (IU question):** *Which hospital do you think is the best for the treatment and recovery time?* (You can select more than one hospital if you think two or more hospitals are equally best) [Answer options: Hospital A/Hospital B/Hospital C/Hospital D, corresponding to the hospitals available in *both* sets]
- **Q3 (Welfare question):** *All things considered, which individual do you think has the best insurance plan?* [Answer options: Mr. Green/Mr. Yellow/Equally good]

Subjects face one of two versions of the survey (between-subject treatments). In version *FreedomIU*, the order of the questions is as presented above, while in version *IUFreedom* the IU question appears first and the freedom question appears second. These orders are kept constant for all set comparisons and the welfare question is always last. The underlying reason for these treatments is two-fold. First, this allows us to test for order effects without increasing participants' cognitive load (as it would be the case, for instance, with a random order of questions for each set comparison). Second, the welfare question is always last as it is framed as an "all things considered" question. While this design choice may make FoC and IU salient in subjects' welfare evaluations, we wanted subjects to consider the welfare question after having compared the two sets in terms of FoC and IU. Doing so allows participants to balance these criteria in their welfare (all-things-considered) evaluations.

[10] We called the two individuals Mr. Green and Mr. Yellow to avoid the influence of individual perceptions and local social norms about names.

[11] All countries in our data have some form of private health insurance plan. Therefore, this formulation is a non-artificial credible reason for opportunity sets to differ across individuals. Still, we do not exclude the possibility that the local organization of hospital care may affect responses.

[12] To avoid staff quality dominating participants' evaluations, we tell them that it excludes doctors and that the hospital choice does not affect Mr. Green or Mr. Yellow's overall health status: "The hospitals available are equivalent in terms of surgery care quality, doctors' skills, etc. Thus, Mr. Green and Mr. Yellow's overall health will not be affected by the choice of hospital for the surgery and the recovering stay".

Table 1
Opportunity set comparisons and theoretical response patterns for cardinality, IU, and range of opportunity rules.

	Set comparisons		Underlying reasons for set comparisons	Theoretical response patterns		
	Set A	Set B		Cardinality	Indirect utility	Range of opportunity
Singleton sets						
s1	{(80, 80)}	{(70, 90)}	Singleton sets without a dominated alternative	∼	A/B/∼	∼
s2	{(81, 92)}	{(80, 75)}	Singleton sets with a dominated alternative	∼	A	A
s3	{(80, 80)}	{(60, 60)}	Singleton sets with a "bad" dominated alternative	∼	A	A
Adding an option						
s4	{(80, 75)}	{(80, 75); (71, 89)}	Adding a non-dominated alternative to a singleton	B	B/∼	B
s5	{(69, 91); (71, 89)}	{(69, 91); (71, 89); (70, 90)}	Adding a non-dominated similar alternative	B	B/∼	B
s6	{(70, 90); (71, 89)}	{(70, 90); (71, 89); (80, 75)}	Adding a non-dominated dissimilar alternative	B	B/∼	B
s7	{(70, 90); (71, 89)}	{(70, 90); (71, 89); (68, 88)}	Adding a "slightly" dominated alternative	B	∼	∼
s8	{(70, 90); (71, 89)}	{(70, 90); (71, 89); (60, 60)}	Adding a dominated "bad" alternative	B	∼	∼
s9	{(70, 90); (71, 89)}	{(70, 90); (71, 89); (20, 30)}	Adding a dominated "very bad" alternative	B	∼	∼
s10	{(80, 75); (95, 95)}	{(80, 75); (95, 95); (71, 89)}	Adding an alternative dominated by a "very good" alternative	B	∼	∼
Trade-off between size and quality						
s11	{(79, 90); (77, 91); (80, 87)}	{(81, 92)}	Trade-off size & quality with a "slightly" dominant alternative	A	B	B
s12	{(80, 80); (69, 91); (71, 89)}	{(81, 92)}	Trade-off size & quality with a dominant alternative	A	B	B
s13	{(70, 70); (50, 71); (72, 65)}	{(81, 92)}	Trade-off size & quality with a "very" dominant alternative	A	B	B
Non-singleton sets of same size						
s14	{(80, 75); (80, 80)}	{(80, 75); (70, 90)}	Non-singleton sets of same size without a dominant alternative	∼	A/B/∼	∼
s15	{(80, 75); (80, 80); (81, 91)}	{(80, 75); (70, 90); (81, 91)}	Non-singleton sets of same size with a dominant alternative	∼	∼	∼

Notes: A pair (x_1, x_2) describes a hospital where x_1 corresponds to staff quality (% rating) and x_2 to comfort quality (% rating). A, B, and ∼ in the columns "theoretical response patterns" refer to the rules' predictions in terms of which set provides more freedom/welfare (set A, set B, or indifference respectively). The indirect utility's theoretical response pattern respects dominance relations; for cases where multiple predictions are shown (s1, s4, s5, s6, and s14), the IU ranking is *a priori* consistent with those predictions, but for each participant it will take only one of the shown predictions depending on the participants' response to which hospital they prefer (Q2).

3.3. The opportunity set comparisons

Participants faced 15 pairwise comparisons of sets presented in a random order. The 15 set comparisons and their rationale are summarized in Table 1. Five main reasons underlie the choice of these set comparisons. First, the theoretical response patterns of different ranking rules are distinctive in these set comparisons. As shown in Table 1, this is particularly the case for size-based rules like cardinality and quality-based rules like IU, which we wanted to clearly distinguish. Second, several of these set comparisons are used as "stress tests" of the ranking rules. For example, set comparisons s7 to s9 are increasingly demanding tests of cardinality's implication that adding an alternative always improves FoC/welfare. Third, these set comparisons were chosen such that subjects would likely face set comparisons for which they considered that there is a "conflict" between FoC and the "best" alternative in the sets. As explained above, these are the only situations where a subject can reveal IvFoC. Fourth, having some set comparisons with a clearly dominant or dominated alternative provides us with an in-built test of attention. For example, if subjects state that in s3 the hospital in set B (60, 60) is better than the hospital in set A (80, 80), then this is a clear mistake. We use this feature of our design to probe our results when we exclude "inattentive" participants. Finally, we focus on small sets with at most three alternatives to limit subjects' cognitive load without restricting our ability to test different theoretical rules, test the intrinsic value of freedom, and bring valuable insights for policy making.

3.4. The spectator position

An important aspect of our design is that we elicit people's attitudes in a *spectator position*. We do this for two key reasons. First, this position is essential to examine whether people rank opportunity sets based on quality-based rules that are prominent in the theoretical literature, such as those that compare sets according to the preferences that could be "reasonably" held by individuals in the relevant context (see the potential preferences family in Section 2.2.3). This could not be achieved using a "stakeholder" design. Second, spectator designs are increasingly used to elicit people's normative attitudes and/or policy preferences. Findings from these studies are seen as relevant input for policy making by shedding light on the political feasibility and democratic desirability of different courses of action. In our context, the spectator position endows participants with similar information to policymakers: our participants, like most policymakers, have incomplete information about people's preferences. Thus, the spectator position offers insights into attitudes from a policy-relevant standpoint, complementing the preferences that would emerge from a stakeholder design.

At the same time, it is important to acknowledge the limitations of this approach. As it can be seen from Table 1, our design allows us to test people's attitudes in situations where it is easy to deduce Mr. Green and Mr. Yellow's preferred option (e.g., s3 and s11-s13). This setup allows us to exclude some alternative explanations for our findings. Specifically, we show that spectators' decisions are *not* driven by their beliefs about Mr. Green and Mr. Yellow's *preferences over alternatives*. Instead, spectators appear

to rely on their own preferences over alternatives when responding to our main questions.[13] However, we cannot exclude the possibility that spectators' beliefs about Mr. Green and Mr. Yellow's *preferences over menus* may influence their menu evaluations. Importantly, the evidence that beliefs about others' preferences over alternatives do not explain our results strongly suggests that spectators' menu evaluations are also unlikely to be driven by their beliefs about others' preferences over menus. Nonetheless, since we cannot fully exclude this possibility and spectators' beliefs about others' preferences over menus may be accurate or not, we interpret our results as primarily capturing spectators' *policy preferences*, i.e., people's preferences from the position of a non-involved party in a policy-relevant domain. This perspective, as argued above, is valuable in its own right and complements the elicitation of stakeholder (incentivized) preferences that could provide valuable insights into how individuals value opportunity sets for themselves. In Section 5, we outline future research directions that could complement our analysis.

3.5. Additional questions

Participants started by stating their gender and age, and at the end of the survey they replied to a short set of questions about their perceived health status, perceived social status, highest level of completed education, occupation, and perceived difficulty of the survey.

3.6. Procedures

Results are based on data from 4902 participants. The mean and median time to complete the survey are respectively 912 (≈ 15 min) and 666 s (≈ 11 min). All these participants finished the survey, responded correctly to an attention question, and passed a speeding check (i.e., took more than 4 min and 15 s to complete the survey). Otherwise, participants were excluded from the data.[14]

We ran the study in ten countries: Brazil (BR), China (CH), Colombia (CO), France (FR), Japan (JA), Netherlands (NL), Portugal (PT), Turkey (TR), United Kingdom (UK), and the United States (US). We collected data in March 2021 using the survey company ODiTY (https://www.areyounet.com), which sent an invitation email to its panel of participants to answer our survey. For completing the survey, participants received "tokens" that they could exchange for money. Each participant received approximately 2€ in each country, with limited variance across countries. To minimize language effects, instructions in English were translated into the local language by professional native speakers, and back-translated to English by another person. Translators were careful to write the instructions in neutral language. The sample is representative of each country in terms of age and gender. Sample characteristics for each country are shown in Appendix C. There is heterogeneity between countries among most observable characteristics, which we control for in our cross-country analysis.

These countries were chosen for two reasons. First, for implementation purposes. In particular, at least one author is fluent in the language of 8 out of 10 of these countries. With the help of two additional colleagues, this allowed us to ensure the quality of the translations for all countries. Second, these countries were chosen for their diversity in terms of social, cultural, and institutional backgrounds. For instance, these countries differ in terms of their dominant religions and political institutions, which may translate into attitudinal differences. The attitudinal differences across these countries are illustrated by the fact that they are spread all over the influential Inglehart-Weltzel world cultural map (see World Values Survey 2022). We therefore believe that this selection of countries provides a somewhat comprehensive (even if clearly incomplete) test for potential cross-country differences.

4. Empirical analysis and results

We structure our analysis as follows. First, we look at how people rank opportunity sets (Section 4.1). We show both aggregate response patterns and our main (Bayes) classifications for freedom and welfare. Second, we investigate if participants attach intrinsic value to freedom of choice (Section 4.2). For these two sections, we use the data from all 4902 subjects. In Section 4.3, we test for cross-country differences.

[13] This aligns with previous findings on individuals' propensity for a *false-consensus bias*, where people tend to overestimate the extent to which others share their preferences (e.g., Engelmann and Strobel 2000, Iriberri and Rey-Biel 2013, Blanco et al. 2014). It also aligns with the notion of *ideals-projective paternalism*, where individuals assume that their own preferences are relevant for guiding others' choices (Ambuehl et al., 2021).

[14] The attention question was randomly presented in the sequence of set comparisons. It consisted of a similar screen to the other set comparisons with an answer option "If you are not a robot please click on this button" (see instructions in Appendix G). The check of 4 min and 15 s was agreed upon with the survey company based on a pilot. Note that since our main research design does not rely on multiple treatments, differential attrition is not an issue. Still, potential differences in attrition, recruitment, and remuneration across countries is one of the reasons why we control for observed characteristics in our cross-country analysis (see Appendix C for attrition data).

Table 2
Subjects' actual response patterns for Freedom and Welfare questions.

	Set comparisons		Freedom question: Which individual do you think has more freedom of choice?			Welfare question: Which individual do you think has the best insurance plan?		
	Set A	Set B	A (% all answers)	B	Same	A	B	Same
Singleton sets								
s1	{(80, 80)}	{(70, 90)}	24.2	10.5	65.3	47.1	14.4	38.5
s2	{(81, 92)}	{(80, 75)}	40.0	3.1	56.9	81.7	3.0	15.3
s3	{(80, 80)}	{(60, 60)}	42.2	2.7	55.0	86.9	2.5	10.6
Adding an option								
s4	{(80, 75)}	{(80, 75), (71, 89)}	6.0	78.6	15.4	7.8	60.0	32.2
s5	{(69, 91), (71, 89)}	{(69, 91), (71, 89), (70, 90)}	6.1	72.9	21.0	8.1	52.2	39.8
s6	{(70, 90), (71, 89)}	{(70, 90), (71, 89), (80, 75)}	7.0	75.3	17.7	9.8	59.5	30.8
s7	{(70, 90), (71, 89)}	{(70, 90), (71, 89), (68, 88)}	8.6	66.9	24.4	14.5	40.7	44.8
s8	{(70, 90), (71, 89)}	{(70, 90), (71, 89), (60, 60)}	10.4	63.9	25.6	18.8	35.6	45.7
s9	{(70, 90), (71, 89)}	{(70, 90), (71, 89), (20, 30)}	14.9	55.2	29.9	27.6	24.6	47.7
s10	{(80, 75), (95, 95)}	{(80, 75), (95, 95), (71, 89)}	8.4	67.6	24.0	10.5	25.2	64.3
Trade-off between size and quality								
s11	{(79, 90), (77, 91), (80, 87)}	{(81, 92)}	72.8	17.0	10.2	28.7	51.8	19.5
s12	{(80, 80), (69, 91), (71, 89)}	{(81, 92)}	73.0	18.0	8.9	27.7	55.4	16.9
s13	{(70, 70), (50, 71), (72, 65)}	{(81, 92)}	68.5	24.0	7.5	15.9	72.5	11.6
Non-singleton sets of same size								
s14	{(80, 75), (80, 80)}	{(80, 75), (70, 90)}	21.6	15.8	62.6	38.6	17.9	43.5
s15	{(80, 75), (80, 80), (81, 91)}	{(80, 75), (70, 90), (81, 91)}	15.4	13.1	71.5	22.4	16.2	61.4

4.1. How people rank opportunity sets

4.1.1. Aggregate response patterns

We start this part of the analysis by showing the aggregate response patterns for the Freedom and Welfare questions. These are summarized in Table 2. Some relevant patterns are already apparent from this table. On the one hand, there are questions for which there is considerable agreement, such as FoC in s4, s6, and s12 as well as welfare in s2, s3, and s13. On the other hand, there is considerable disagreement in others, such as FoC in s2 and s3 as well as welfare in s1, s7-s9, and s14. In addition, it is noticeable that responses to the Freedom and Welfare questions differ significantly. Note that these differences are in the expected direction. For example, most (though not all) people consider that A offers more FoC than B in s11 to s13, while most (though far from all) people consider that B offers more overall welfare than A in these set comparisons.

The patterns for each situation in Table 2 strongly suggest that answers are not random. Answers across situations for the same question (either Freedom or Welfare) also support non-randomness. For example, as expected, it is more often the case in s7 than in s8 and in s8 than in s9 that set B is considered to have more FoC than set A. Noise — which is common in similar studies — also seems moderate in our setting. For example, in s3 there are only 2.5% (10.6%) that state that B provides more (same) welfare than A when clearly set A dominates set B.[15]

4.1.2. Main classifications

Each ranking rule offers a prediction in terms of which set provides more freedom/welfare for each set comparison. The *theoretical response patterns* of the most prominent rule of each family are summarized in Table 1 (see Table A.2 in Appendix A for the theoretical response patterns of all rules).

We can then compare the rules' theoretical response patterns to the subjects' actual response patterns in order to "classify" participants according to the rule they implicitly employ, while incorporating a certain probability of noise in the subjects' choices. We do this for both questions (Freedom and Welfare). Here we follow (Ambuehl and Bernheim, 2021) who apply the methods from (Hastie et al., 2001) and Costa-Gomes et al. (2001) to perform a Bayes classification exercise similar to ours. We assign to each subject a rule among the 23 plausible theoretical rules described above and an error probability. To do this, we use the *Maximum a posteriori* rule, i.e., we maximize the posterior probability of following a rule with a given error and conditional on the subjects' response pattern. Formally, denote by c_i the response pattern of subject i, i.e., the vector of answers that subject i gives to the $S = 15$ binary set comparisons. When following a certain rule r_j (with $j \in \{1, \dots, J\}$; $J = 23$ in our analysis), a subject has

[15] In the IU question (Q2), we ask participants which hospital they think is best for *treatment and recovery time*. *Treatment and recovery time* refers to the total time spent in hospital, for which hospitals differ in terms of staff quality and comfort quality. However, as we stated that "Mr. Green and Mr. Yellow's overall health will not be affected by the choice of hospital for the surgery and the recovering stay" (fn. 12), there was a potential risk that participants might perceive all hospitals as equivalent in terms of *treatment and recovery time*. Yet, in set comparisons where we do *not* expect indifference in the IU question (i.e., s1 to s3 and s11 to s13), we observe an average of 7.03% indifference reports, ranging from 2.12% in s3 to 11.65% in s1. This finding strongly suggests that participants, as intended, recognized differences between hospitals in this dimension.

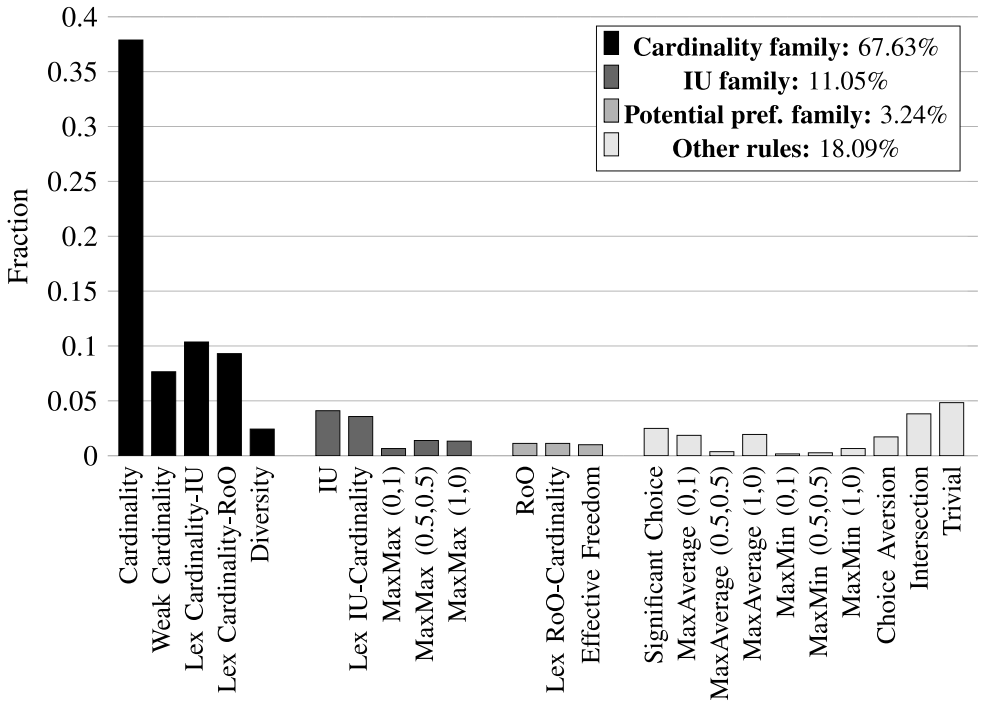

Fig. 1. Bayesian classification for Freedom question.

a certain probability of making an error, denoted by ε_j. More precisely, she follows the rule with probability $1-\varepsilon_j$ and uniformly randomizes over the three possible choices (A, B and $\sim$) with probability ε_j. We also assume that ε_j are distributed uniformly over $[0,1]$ and independent across profiles. Denote by $P(r_j,\varepsilon_j|c_i)$ the posterior probability to follow a rule r_j with error probability ε_j conditional on the response vector c_i. We assign a subject to rule r_j if and only if:

$$P(r_j,\varepsilon_j^*|c_i) > P(r_k,\varepsilon_k^*|c_i) \quad \forall \quad k \in \{1,\dots,J\} \setminus j \tag{1}$$

where ε_j^* is the probability that maximizes $P(r_j,\varepsilon_j|c_i)$, i.e., ε_j^* is the probability that maximizes the likelihood that a subject i with response vector c_i follows the rule r_j. When more than one rule maximizes the posterior probability, we assign the subject to one of these rules at random. See Appendix B for further technical details.

Classification results for freedom

The main classification results for the Freedom question are displayed in Fig. 1. The cardinality rule is the most often implicitly used by subjects: 38% of spectators make choices that are consistent with this rule. In addition, the weak cardinality rule gathers 8% of subjects, and 20% have response patterns consistent with lexicographic rules that give priority to the menu size. All in all, a large majority (68%) of participants rank sets in terms of FoC according to cardinality family rules. On the contrary, quality-based rules are seldom used. The IU family accounts for 11% of participants, while the potential preferences family accounts for only 3% of subjects.

How good is the fit between our classification and the subjects' choices? To see this, we look at the average error (ε^*) for each ranking rule, i.e., the probability that the average subject chooses at random. The first row of Table 3 reports the average error for the main rules and families for the Freedom question. The total average error is 0.21 (i.e., on average subjects choose at random in around 3 out of 15 set comparisons). Note that this error is considerably lower for the most prevalent rule: subjects assigned to the cardinality rule make random choices in only 1.5 set comparisons on average. Overall, the fit is very good for the 46% who follow cardinality and weak cardinality, good for a significant proportion of other subjects (such as the 24% assigned to Lex Cardinality-RoO, Lex Cardinality-IU, and IU), and less good for the least prevalent rules.

Table 4 shows the choices for the different set comparisons for participants that follow the cardinality rule. This data allows us to identify where subjects depart from this rule. Three insights come out of it. First, the goodness-of-fit is high for all questions, and for 12 out of 15 questions the consistency is larger than 92%. Second, not surprisingly, it seems that participants who follow cardinality are most likely to depart from it when we add an option of "poor" quality (set comparison s9). Third, deviations are also more common when these subjects compare non-singleton sets with the same size (see s14 and s15). While some of these deviations may be due to random mistakes, these patterns suggest that a small number of subjects assigned to cardinality are sensitive to quality considerations in these three set comparisons.

Table 3
Goodness-of-Fit of the Bayesian classifier.

Question	**Average**	**Cardinality family**	Cardinality	Weak cardinality	Lex Cardinality-IU	Lex Cardinality-RoO	Diversity	**IU family**	Indirect Utility	Lex IU-Cardinality	MaxMax (0,1)	MaxMax (0.5,0.5)	MaxMax (1,0)	**Potential Pref. family**	**Other rules**
Freedom	0.21	0.15	0.10	0.08	0.31	0.24	0.29	0.33	0.23	0.38	0.43	0.39	0.39	0.36	0.34
Welfare	0.30	0.33	0.26	0.23	0.35	0.33	0.65	0.28	0.21	0.32	0.33	0.28	0.32	0.27	0.33

Notes: This table reports the average error (ε^*) for main ranking rules and families, i.e., the probability that the average subject chooses at random. "Cardinality family", "IU family", and "Potential Pref. family" report the average error for the rules in the respective families, while "Other rules" reports the average error for the theoretical rules that do not belong to these families.

Table 4
Choices made by subjects classified as Cardinality (Freedom question).

Comparisons:	s1	s2	s3	s4	s5	s6	s7	s8	s9	s10	s11	s12	s13	s14	s15
Prediction:	~	~	~	B	B	B	B	B	B	B	A	A	A	~	~
Choice freq.:															
A	5	5	5	2	1	2	2	3	4	2	98	98	96	7	7
~	92	94	94	1	0	3	3	4	11	3	1	1	1	85	87
B	3	1	1	97	99	95	95	93	85	95	1	1	3	8	6

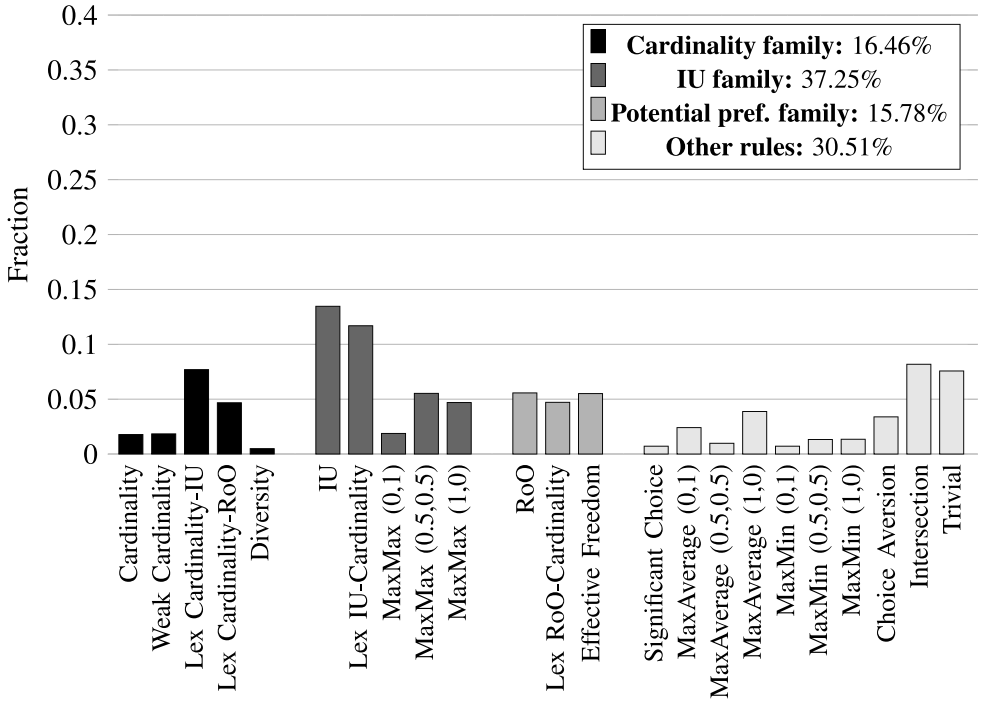

Fig. 2. Bayesian classification for Welfare question.

Classification results for welfare

The main classification results for the Welfare question are displayed in Fig. 2. As can be seen from the figure, the most prevalent rule is IU. However, only 14% of subjects respond as if following this rule. Despite the heterogeneity of the classification, two patterns are apparent. First, a significant number of subjects follow quality-based rules. In particular, 37% of participants follow IU family rules (14% IU, 12% Lex IU-Cardinality, and 12% MaxMax) and 16% are consistent with potential preferences family rules. Second, cardinality-based rules are seldom used to evaluate the welfare of sets, with only 16% of subjects following a cardinality family rule for this question (with most of these assigned to Lex Cardinality-IU and Lex Cardinality-RoO).

It is also worth noting that these classification results have less goodness-of-fit than the ones for the Freedom question. As reported in the second column of Table 3, the average error is 0.30 for the Welfare question. This error is lower for the most prevalent rule

Table 5
% of choices made by participants classified as IU consistent with IU.

Comparisons:	s1	s2	s3	s4	s5	s6	s7	s8	s9	s10	s11	s12	s13	s14	s15
% of choices:	91	93	95	83	85	86	84	82	81	88	78	81	95	87	74

(0.21 for IU), and moderate for other quality-based rules belonging to the IU family (0.28 on average) and the potential preferences family (0.27 on average). Overall, it is apparent that subjects are more idiosyncratic in their welfare evaluations than in their freedom evaluations.

Finally, Table 5 reports the consistency between IU's rule predictions and the choices made by subjects classified as IU.[16] We can see that the goodness-of-fit is high for most questions. For the questions for which the consistency is lower, the most interesting pattern is from s9, where subjects assigned to IU who contradict it tend to prefer the smaller set (131 out of 158). This suggests that these subjects consider that a "very bad" alternative is detrimental to the quality of the set. This trend is aligned with our aggregate result that 28% of our sample prefers the smaller set in s9 (see Table 2). This could be explained by a concern that decision-makers would commit mistakes when choosing (see Bartling et al. 2023 for related evidence).

Robustness checks and alternative explanations

Several checks reported in Appendix D provide further confidence to our main classifications. First, our results are robust when we perform checks for attention. The classifications are very similar when we remove subjects that make clear mistakes in their comparisons of sets, such as stating that (60, 60) is better than (80, 80) (Appendix D.1).

Second, results are also very similar when we perform our analysis per quartiles of total response times (Appendix D.2). The latter analysis also shows that cardinality is *not* being used as a heuristic to quickly respond to the questions: contrary to this hypothesis, the fastest 25% of participants are significantly less likely to follow cardinality in the Freedom question. In the same direction, subjects assigned to the cardinality family are statistically significantly slower than others.[17]

Third, as reported in Appendix D.3, our results are robust to the order of the questions (versions *FreedomIU* and *IUFreedom*). Consistent with order effects and the relevance of salience, we find that subjects in the freedom classification are statistically significantly less likely to be assigned to the cardinality rule in *IUFreedom* than *FreedomIU* (i.e., when they are first asked the IU question). However, this effect is small and does not change our main results (42% and 34% are assigned to cardinality in *FreedomIU* and *IUFreedom* respectively). In addition, the overall classifications of both freedom and welfare are otherwise very similar across the two versions.

Fourth, our classifications using subjects' responses do much better than a classification using random answers (Appendix D.4). In particular, with random answers, no rule gathers more than 9% of (artificial) subjects and the average error is 0.70 (compared to 0.21 and 0.30 for our freedom and welfare classifications respectively).

Fifth, a cluster analysis using a k-modes procedure to infer rules from the data (Chaturvedi et al., 2001) produces results that closely align with those obtained through the Bayesian classifier used to fit rules to the data (see Appendix D.5 for results). In particular, all our main findings hold. This analysis also reinforces the results reported above indicating that a small subset of participants exhibits choice aversion in certain set comparisons (specifically s7 to s9). In addition, this analysis suggests that, when evaluating the welfare of sets, approximately 15% of subjects implicitly use the following variant of the IU rule: Set C is overall better than D when it contains a hospital that clearly dominates the best hospital in D (as in s2, s3, and s13), otherwise both sets provide the same overall welfare.

Sixth, our findings remain robust when we impose, *ex post*, a maximal bound on the error probability ε^*. Specifically, we reanalyzed our main classification using maximal bounds set at 0.3, 0.5 and 0.7 (see Appendix D.6 for detailed results). As shown in the appendix, these analyses are again consistent with our main conclusions. Among participants who can be assigned to a rule under these maximal error bounds, cardinality-based rules dominate their freedom evaluations, while IU-based rules are the most prevalent in their welfare evaluations — though significant heterogeneity remains in the latter. Additionally, imposing these bounds strengthens the cluster analysis' finding that close variants of our rules may play a role in explaining the response patterns of several participants when evaluating the welfare of opportunity sets.

Finally, our main results hold when we restrict the analysis to set comparisons where it is easy to deduce Mr. Green and Mr. Yellow's preferences over alternatives (s2, s3, s7 to s13 and s15; see Appendix D.7 for results). This shows that our results are not driven by the uncertainty about others' preferences. It also demonstrates that even if participants had incomplete or intransitive preferences over some hospitals, this would not affect our main results.

[16] We cannot report a table equivalent to Table 4 because the predictions of IU depend on subjects' preferences over hospitals (Q2).

[17] Average and median times to complete the survey are respectively 944 and 700 s for subjects assigned to the cardinality family and 845 and 586 s for others ($p = 0.0085$ for a Student test for difference between means and $p < 0.001$ for a Wilcoxon-Mann–Whitney test).

Table 6
IvFoC: Main results.

	Relevant cases		IvFoC ($C \succ_i^W D$ in relevant cases)					
	Mean (out of 15)	% subjects > 0	Mean	% of subjects s.t. IvFoC revelations = 0	≥ 1	≥ 2	≥ 3	≥ 4
$C \succ_i^F D$ and								
$\max(D)R_i \max(C)$	7.15	96.43	3.42	15.12	84.88	70.60	57.67	45.84
$\max(D)I_i \max(C)$	5.13	95.19	2.87	17.50	82.50	66.67	51.96	38.43
$\max(D)P_i \max(C)$	2.02	80.74	0.54	66.10	33.90	14.46	4.92	0.78

Notes: The first column shows the mean number of responses per subject (in a total of 15 set comparisons) such that C (either set A or B) is considered to have more freedom than D (A if $C = B$ or B if $C = A$) and $\max(D)R_i \max(C)$ (first row), $\max(D)I_i \max(C)$ (second row), or $\max(D)P_i \max(C)$ (third row). The second column shows the percentage of participants for which the number of relevant cases is at least one. The third column shows the mean number of responses per subject such that $C \succ_i^W D$ in the relevant cases (i.e., when $C \succ_i^F D$ and $\max(D)R_i \max(C)/\max(D)I_i \max(C)/\max(D)P_i \max(C)$). The last five columns present the percentage of subjects who reveal IvFoC among the participants for which the number of relevant cases is at least one.

4.2. The intrinsic value of freedom

Table 6 reports our main results for the intrinsic value of freedom. The first two columns suggest that there is a significant number of "relevant cases", i.e., set comparisons for which subjects express that there is a "conflict" between FoC and the "best" alternative. These are the only relevant set comparisons since, by construction, a spectator cannot reveal attaching intrinsic value to freedom when set A has more FoC and it has a better alternative than all available in set B. In a total of 15 set comparisons, the average number of situations with a "conflict" between FoC and the best alternative is close to 50% (7.15 out of 15) and almost all subjects (96%) are presented with at least one set comparison for which they answered that $C \succ_i^F D$ and $\max(D)R_i \max(C)$. From these, the majority are cases in which subjects are indifferent between the best element of C and D $[\max(D)I_i \max(C)]$, but there is a significant number of cases where subjects strictly prefer the best element of D to that of C $[\max(D)P_i \max(C)]$. The latter cases present a clear conflict between FoC and the best alternative, while the former present a "mild" conflict in the sense that a subject can still contradict standard economic theory by revealing that he/she attaches intrinsic value to freedom.

How many spectators reveal that they attach intrinsic value to freedom *in these relevant cases*? The results are striking. On average, 48% of responses reveal attaching intrinsic value to freedom in these situations (3.42 out of 7.15). Moreover, 85% of spectators reveal that they value freedom intrinsically at least once, with a majority (58%) doing so at least three times. A large proportion of these IvFoC revelations occur in cases where participants are indifferent between the best alternative in the two sets (see the second row of Table 6). Still, a significant number of subjects express that a larger set provides greater overall welfare than a smaller set even though the larger set has a "worse" (dominated) top alternative. In fact, 34% of spectators reveal attaching such a "high" intrinsic value to freedom at least once, despite having, on average, only two opportunities to do so (see the third row of the first and fifth columns of Table 6). As shown below, these results cannot be attributed to response error or alternative explanations such as uncertainty about others' preferences. Overall, these findings challenge the conventional view that freedom holds only instrumental value.

Table 7 reports the proportion of spectators who reveal that they attach IvFoC for each set comparison, providing further insights into when and why individuals exhibit this attitude in policy-relevant contexts. Results can be summarized as follows. First, it is difficult to rationalize responses that attach IvFoC when comparing singleton sets (s1 to s3). However, as shown in the table, the number of participants doing so is very small (41.37%*278=115 for s1, 69 for s2, and 53 for s3). Second, for set comparisons where we "add an option" (s4 to s10), the IvFoC is driven by cases where participants are indifferent between the best alternative of both sets (i.e., s4 to s10 in column ii). This is not surprising since, in set comparisons like s7 to s10, the "best" alternative is present in both sets. Overall, IvFoC ranges from 33% to 69% of relevant cases for these set comparisons (s10 and s4 respectively, as shown in column ii), and its prevalence appears to depend on both the quality of the best alternative *and* the nature of the "added" option. Third, set comparisons involving a "trade-off between size and quality" (s11 to s13) are primarily responsible for the earlier finding that a significant proportion of spectators attach such high intrinsic value to freedom that they judge a larger set to be overall better even if it contains a "worse" top alternative (see s11 to s13 in column iii). Here again, the quality of the alternatives seems to matter. Finally, fewer participants perceive a "conflict" between freedom and the best alternative when comparing non-singletons of the same size (s14 and s15). Nonetheless, even in these cases, a non-negligible number of participants seem to evaluate the overall welfare of the sets based on more information than just their top alternatives.

Robustness checks and alternative explanations

Several checks provide further support to these findings (see Appendix E). First, as with our classification results, these findings are very similar when we remove subjects that made clear mistakes in their comparisons of sets (Appendix E.1), as well as when we perform our analysis per quartiles of total response times (Appendix E.2). The IvFoC is also robust to the order of the questions (versions *FreedomIU* and *IUFreedom*). Once again, even though there are statistically significant order effects, the overall results are strikingly similar (Appendix E.3).

Our results are also robust to four potential alternative explanations. First, a reader might wonder whether response errors could rationalize the intrinsic value of freedom observed in our study. In Appendix E.4, we use dominance relations of alternatives across

Table 7
IvFoC: Results per set comparison.

	Set comparisons		IvFoC ($C \succ_i^W D$ in relevant cases)					
	A	B	(i) $C \succ_i^F D$ and $\max(D)R_i \max(C)$		(ii) $C \succ_i^F D$ and $\max(D)I_i \max(C)$		(iii) $C \succ_i^F D$ and $\max(D)P_i \max(C)$	
			% IvFoC	# relevant cases	% IvFoC	# relevant cases	% IvFoC	# relevant cases
Singleton sets								
s1	{(80, 80)}	{(70, 90)}	41.37	278	60.47	43	37.87	235
s2	{(81, 92)}	{(80, 75)}	48.59	142	60.00	35	44.86	107
s3	{(80, 80)}	{(60, 60)}	41.73	127	66.67	21	36.79	106
Adding an option								
s4	{(80, 75)}	{(80, 75), (71, 89)}	69.06	3,038	69.23	2,896	65.49	142
s5	{(69, 91), (71, 89)}	{(69, 91), (71, 89), (70, 90)}	62.51	2,942	62.89	2,881	44.26	61
s6	{(70, 90), (71, 89)}	{(70, 90), (71, 89), (80, 75)}	69.04	2,571	68.69	2,427	75.00	144
s7	{(70, 90), (71, 89)}	{(70, 90), (71, 89), (68, 88)}	58.66	3,643	58.75	3,634	22.22	9
s8	{(70, 90), (71, 89)}	{(70, 90), (71, 89), (60, 60)}	53.47	3,585	53.48	3,573	50.00	12
s9	{(70, 90), (71, 89)}	{(70, 90), (71, 89), (20, 30)}	48.78	3,403	48.78	3,393	50.00	10
s10	{(80, 75), (95, 95)}	{(80, 75), (95, 95), (71, 89)}	34.26	3,657	33.44	3,595	82.26	62
Trade-off between size and quality								
s11	{(79, 90), (77, 91), (80, 87)}	{(81, 92)}	30.88	3,177	48.02	429	28.20	2,748
s12	{(80, 80), (69, 91), (71, 89)}	{(81, 92)}	29.02	3,218	50.00	370	26.30	2,848
s13	{(70, 70), (50, 71), (72, 65)}	{(81, 92)}	18.18	3,207	49.73	185	16.25	3,022
Non-singleton sets of same size								
s14	{(80, 75), (80, 80)}	{(80, 75), (70, 90)}	60.33	842	71.81	518	41.98	324
s15	{(80, 75), (80, 80), (81, 91)}	{(80, 75), (70, 90), (81, 91)}	67.85	1,238	68.46	1,151	59.77	87
All			47.80	35,068	56.03	25,151	26.93	9,917

Notes: This table reports the % of participants that prefer C (either set A or B) to D (A if $C = B$ or B if $C = A$) given that they consider that C provides more freedom than D while the preferred element(s) of D is (i) weakly preferred, (ii) indifferent to, or (iii) strictly preferred to that of C. "# relevant cases" denotes the number of responses/participants per set comparison such that $C \succ_i^F D$ and $\max(D)R_i \max(C)/\max(D)I_i \max(C)/\max(D)P_i \max(C)$.

sets to test if participants consistently reveal IvFoc in set comparisons where consistency is expected.[18] Specifically, we expect that subjects who reveal IvFoC in s9 (s13) should also reveal it in s8 and s7 (s12 and s11), and those who reveal it in s8 (s12) should do so in s7 (s11). Consistency levels are above 73.5% in all cases, with one exception at 65.5% (s12 to s11). Additionally, we show that response errors are more likely to lead to an underestimation, rather than an overestimation, of the observed level of intrinsic value of freedom. In a nutshell, we show that cases where subjects prone to response errors fail to reveal IvFoC despite attaching intrinsic value to freedom, are expected to be more common than cases where participants reveal IvFoC without actually attaching intrinsic value to freedom. Taken together, these results demonstrate that response errors cannot explain the observed intrinsic value of freedom in our study.

Second, *preferences for flexibility* (i.e., a preference for larger sets to better tailor for multiple potential future preferences), could, at least in principle, explain the intrinsic value of freedom in some of our set comparisons (see Arlegi and Nieto 2001a, 2001b for a general argument). This could be problematic, since preferences for flexibility are an instrumental concern about the benefits that FoC can entail in the future. However, in our setting, preferences for flexibility could only explain IvFoC in set comparisons where a non-dominated alternative is added to the set (s4 to s6). They cannot rationalize the behavior of a participant who reveals IvFoC in set comparisons where dominated alternatives are added (s7 to s10). They can neither explain revelations of IvFoC in other set comparisons such as s11 to s13 and s15. However, Table 7 shows that IvFoC is revealed in 40% of the relevant cases in the latter set comparisons (s7 to s13 and 15). It follows that preferences for flexibility cannot be the underlying reason for the intrinsic value of freedom in our setting.[19]

Third, and relatedly, attaching intrinsic value to freedom in our setting is not driven by uncertainty regarding the preferences of others. As shown in Table 7, IvFoC holds in set comparisons where it is easy to deduce Mr. Green and Mr. Yellow's preferences over alternatives *and* these preferences should be equivalent to those of participants (s7 to s13 and s15).

Finally, participants could potentially compare opportunity sets based on attributes that are not explicitly mentioned in our vignette. For example, they could consider potential prices of the insurance plans when responding to the Welfare question. We deliberately used insurance plans to enhance the "ecological validity" of our vignette (Brunswik, 1955), as it offers a credible reason for differences in the opportunity sets available to individuals. Alternative formulations would seem less realistic, and mentioning prices could have had the unintended consequence of making it salient. A related concern might be that participants could perceive

[18] We do not employ a similar method to the one used in Section 4.1 to account for response errors because the IvFoC revelation principle introduced in Section 2.1 does not provide a theoretical response pattern across the 15 set comparisons. This means that, among other things, a subject can reveal IvFoC in a single set comparison without contradicting our IvFoC revelation principle. Therefore, an analysis similar to Section 4.1 is not feasible.

[19] Differences in levels between s4 to s6 and s7 to s10 are not directly comparable since the alternatives added are not of similar quality. We therefore abstain from making an inference if preferences for flexibility can or cannot explain a small percentage of observed IvFoC.

a larger number of hospitals in an opportunity set as increasing the likelihood of having a hospital closer to one's address. If participants were making such assumptions, they could reveal IvFoC (in the above sense) for instrumental reasons, which would go against our interpretation. To see this, consider situation s4, where menu A offers one hospital (80, 75) and menu B offers two hospitals (80, 75) and (71, 89). Suppose a participant prefers (80, 75) to (71, 89), considers that menu B offers more freedom than menu A, and believes that menu B provides greater overall welfare than menu A. This would be a revelation of IvFoC. The alternative explanation based on proximity could only explain this pattern of responses if the participant would choose the closer hospital (71, 89) over the farther hospital (80, 75), even though the latter is the preferred one for treatment and recovery time. While this reasoning could rationalize IvFoC in certain set comparisons, it is unlikely to do so in others. For instance, it is improbable that having the hospital $(20, 30)$ in s9 could be valued for instrumental (proximity) reasons. Given that we observe nearly 50% of IvFoC revelations in the relevant cases of s9, this suggests that an intrinsic preference for choice, rather than instrumental factors, is driving most of the observed intrinsic value of freedom.

4.3. Cross-country results

Are these results dependent on the social, cultural, and institutional background of participants? To answer this question, we compare our main results across 10 countries with very different social, cultural, and institutional backgrounds. Since our subject pools may differ across countries due to differences in recruitment and other socio-economic characteristics, we control for observed characteristics in our comparison across countries. In accordance, we present results for the "average participant" (i.e., a participant whose characteristics are averaged over all countries).

4.3.1. Main classifications

To test if people from different countries rank sets differently, we first estimate a multinomial logit regression where the dependent variable is the family of rules to which the rule assigned to the respondent belongs (cardinality family, IU family, potential references family, and others), and the independent variables are dummies for countries and observed characteristics (age and binary variables for gender, health status, social category, education level and occupation; see Table C.1 in Appendix C for details). We then estimate, for each country, the predicted probability of the average participant to be assigned to a given family as if he/she is from that country.

Fig. 3 reports, for each country, the predicted probability that an average participant is assigned to a rule from the cardinality family, IU family, or potential preferences family. For the Freedom question (left panel), the cardinality family dominates across all countries, with between 57% (Japan and China) and 81% (Portugal) of participants assigned to this family. Some country differences are statistically significant (see Table F.2 in Appendix F). By contrast, the IU and potential preferences families account together for less than 20% of participants in all countries. Therefore, despite visible cross-country differences, the overarching pattern is consistent across countries: the majority of subjects rank freedom of choice as if following size-based rules, while only a small minority follows quality-based rules when evaluating the freedom of menus.

For the Welfare question (right panel), the IU family is the most prevalent across countries (between 32% in China and 42% in Japan), while the potential preferences family gathers between 12% (China) and 19% (UK) of subjects across countries. Most of these differences are small and most (though not all) are statistically insignificant at 5% (see Table F.3 in Appendix F). On the other hand, there are between 8% (Japan) and 24% (Turkey) participants across countries that rank the welfare of sets following size-based rules. Again, despite some visible differences, the message is robust across countries: A significant proportion of participants use IU-based rules and other quality-based rules to judge the welfare of sets, but there is considerable heterogeneity in how people evaluate the welfare of sets.

4.3.2. Intrinsic value of freedom

Fig. 4 reports, for each country, the predicted average number of times that the average participant reveals IvFoC and the predicted probability that an average participant exhibits IvFoC at least once. As shown in the figure, the overall picture is very similar across countries. In all countries, the predicted average number of times that the average participant reveals IvFoC is between 2.89 (France) and 3.98 (Turkey) (left panel). Although many of these differences are statistically significant, the magnitude of these differences is rather small (see Table F.4 in Appendix F). Results are again very similar when we look at the predicted probability that an average participant reveals IvFoC at least once (right panel). Even though some of these differences are statistically significant, the effect size does not question our main findings (see Table F.5 in Appendix F).[20] Overall, these results demonstrate that attaching intrinsic value to freedom of choice in our setting is a robust cross-cultural phenomenon independent of participants' social, cultural, or institutional background.

[20] In Appendix F, we show that differences across countries are mainly driven by cases where participants are indifferent between the top alternatives of both sets.

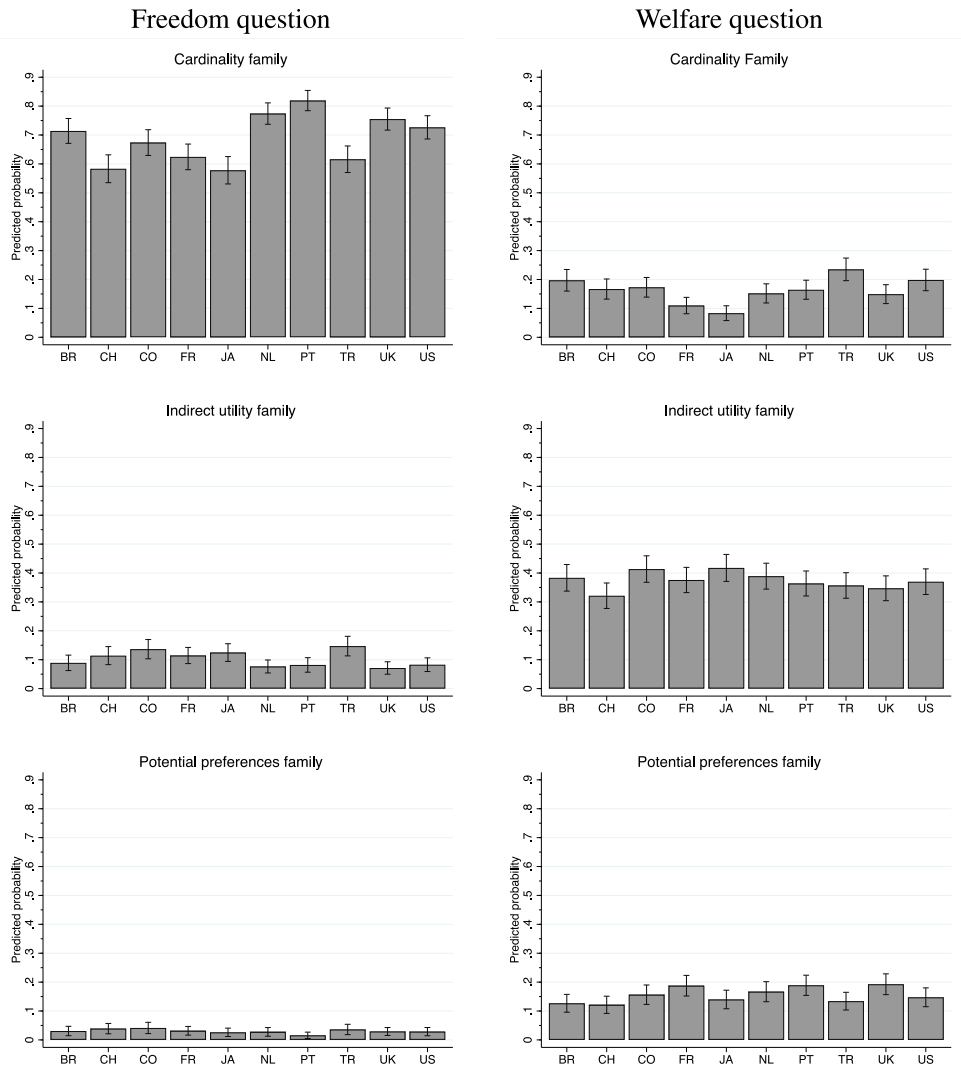

Fig. 3. Classifications per country.
Notes: These figures report the predicted probabilities of an average participant to be assigned to the cardinality family, IU family, and potential preferences family for the Freedom question (left panel) and the Welfare question (right panel). These estimates are based on multinomial logit regressions controlling for observed characteristics (see Table F.1 in Appendix F). Lines indicate 95% confidence interval.

5. Concluding remarks

In economics, it is standard practice to evaluate social states based on individual preferences. "Freedom" is often regarded as an important value, but only for instrumental reasons: giving individuals freedom of choice enables them to select the best outcomes according to their individual preferences, which are typically unknown to policymakers. This near-exclusive emphasis on outcomes, evaluated in terms of preferences, contrasts with public debates in many countries where "freedom" is frequently invoked as a value in its own right without reference to preferences. Our paper sheds new light on this topic. We employ a novel survey-based "spectator" research design to investigate two key questions: (i) whether people attach intrinsic value to freedom of choice in a policy-relevant context, and (ii) which theoretical rules they implicitly use when ranking opportunity sets in terms of both freedom of choice and overall welfare. We do this for a total of 4902 participants across 10 countries with distinct social, cultural, and institutional backgrounds. Surprisingly, we find that a majority of spectators *reveal* that they attach intrinsic value to freedom in the context of hospital choice. Additionally, most participants rely on size-based (cardinality) rules to rank menus of options in terms of freedom, while there is considerable heterogeneity in the theoretical rules they implicitly employ when ranking opportunity sets in terms of welfare. These findings are remarkably consistent across countries.

Results in positive welfare economics can help bridge the gap between people's intuitions and theoretical principles (see, e.g., Gaertner and Schokkaert 2012, Ambuehl and Bernheim 2021). Our findings contribute valuable insights to this discussion

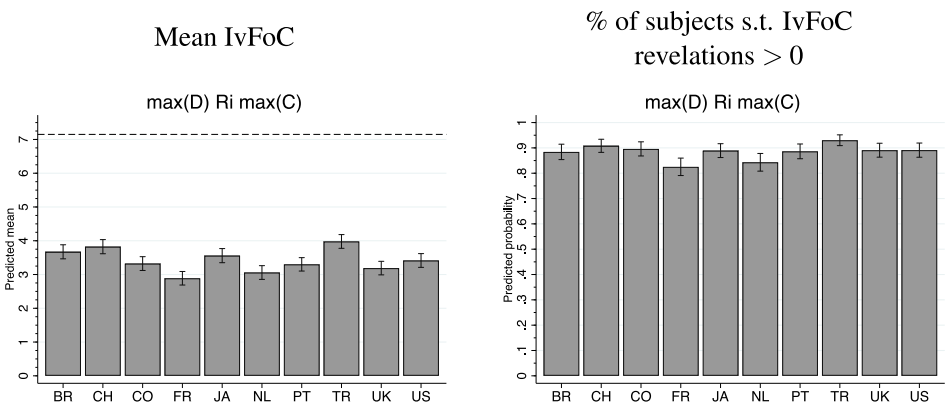

Fig. 4. Intrinsic value of freedom per country.
Notes: The left panel reports the predicted average value of the number of responses such that C (either set A or B) is considered to provide more welfare than D (A if $C = B$ or B if $C = A$), when C is considered to have more freedom than D and $\max(D)R_i\max(C)$, using an OLS regression controlling for the number of relevant cases and observed characteristics. The right panel reports the predicted probability of this number being positive, using an OLS regression controlling for the same variables. See Table F.6 in Appendix F for underlying regressions. Lines indicate 95% confidence interval.

by revealing how individuals think about the freedom and welfare of opportunity sets. On the one hand, our results support some theoretical perspectives. For instance, some authors argue that counting the number of alternatives is "a natural way of measuring freedom in the absence of information about the agent's preferences" (Foster 2011, p. 20). Consistent with this, we find that most people reason about freedom of choice in this way. On the other hand, our results challenge other theoretical views. Concepts like potential preferences and the "eligibility" of options, while widely endorsed in the FoC literature, do not seem to influence how most people perceive which options increase freedom of choice. Moreover, contrary to the traditional view in economics, welfare evaluations of opportunity sets are multidimensional and vary significantly across people. This suggests that while evaluating sets based on their indirect utility may be a sensible and parsimonious simplification in certain contexts, it will overlook important welfare-enhancing features of opportunity sets that people care about (see Benjamin et al. 2014 for related evidence). Overall, these findings offer valuable insights to both positive and normative theories of how people rank (or should rank) opportunity sets.

From a policy perspective, many authors have argued that public policy should *not* focus on satisfying preferences, but on providing opportunities for individuals to achieve their own ends, whatever those ends might be (e.g., Rawls 1971, Roemer 1998, Sugden 2004). Even if one does not endorse this view, people's attitudes toward menus can still *complement* information about their preferences over alternatives as inputs to public policy. Our findings can inform policymakers on how to compare different opportunity sets while taking these attitudes into account. For example, our findings highlight that people value the expansion of small opportunity sets, even when this does not result in different choices being made. This insight supports policy reforms that increase choice in small sets, as such reforms can be beneficial even if individuals do not modify their behavior. A notable example is the UK's National Health Service, where the number of hospitals that patients can choose from has changed over time — ranging from one to five — with the most recent reform introduced by the UK Government in 2023.

Our study opens several avenues for future research. One important question is the extent to which the intrinsic value of freedom influences stakeholder decision-making. Are people willing to pay for additional freedom even when it does not provide better options? To explore this, one could apply our novel elicitation method — the IvFoC revelation principle — in contexts where decision-makers face "real" stakes, such as consumption decisions in incentivized experiments. It would also be important to study how people rank opportunity sets with larger numbers of options, investigating the relationship between freedom of choice and the well-documented difficulties of decision-making in large menus (see, e.g., Iyengar and Lepper 2000, Chernev et al. 2015). Another important avenue for future research would be the relationship between spectators' preferences over menus and their beliefs about others' preferences over menus. For example, do individuals hold accurate beliefs about others' menu preferences, or are these beliefs systematically biased? In addition, it would be relevant to understand how people's rankings of opportunity sets relate to other attitudes, such as paternalism, and how these rankings might influence behavior in different contexts, such as voting.

Future research should also explore whether our findings on spectators' policy preferences extend beyond hospital choice. The health sector is a critical area of study, as it represents a significant share of GDP in OECD countries and there are ongoing debates about expanding personal choice in healthcare in countries like France, Portugal, and the UK. While the abstract nature of our survey questions may facilitate generalizability to similar contexts, such as school choice, applying our method to other domains introduces challenges. For instance, infrequent and complex decisions, such as those about housing, are likely to involve preferences that are formed during the decision-making process itself. Similarly, studying hard ethical choices would be an important and challenging setting, where freedom of choice may conflict with a willingness to avoid responsibility. Should surgeons have greater freedom in allocating scarce medical resources or making battlefield triage decisions? Finally, another critical issue is the relationship between externalities and freedom of choice, as exemplified by the interplay between consumption decisions and climate change. These important questions remain for future research.

Data and replication package

The replication package for this paper is available at https://doi.org/10.5281/zenodo.15057699.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We are grateful to Sandro Ambuehl, Alexander Cappelen, Olivier L'Haridon, Yannick L'Horty, Stephane Luchini, Séverine Toussaert, Michael Vlassopoulos, and participants of workshops in Lyon, Southampton, Osaka, Paris (SCW meeting), Aussois (TEPP winter seminar), and LSE (International Inequalities Institute) for valuable discussions, comments and suggestions. We are also grateful to professional and friend translators and Sizar Mani from Odity for data collection. This paper forms part of the research project "ValFree" (The Value of Choice, grant No. ANR-16-CE41-0002-01) of the French National Agency for Research whose financial support is gratefully acknowledged. Hanaki gratefully acknowledges financial support from grants-in-aid for scientific research (KAKENHI, grant numbers: 18K19954, 20H05631) from the Japan Society for the Promotion of Science (JSPS) as well as support from the Joint Usage/Research Center at ISER, Osaka University.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.euroecorev.2025.105022.

Data availability

Replication package for "Freedom counts: Cross-country Empirical Evidence" (Original data) (ZENODO)

References

Ackfeld, V., Ockenfels, A., 2021. Do people intervene to make others behave prosocially? Games Econom. Behav. 128, 58–72.
Almås, I., Cappelen, A., Tungodden, B., 2020. Cutthroat capitalism versus cuddly socialism: Are Americans more meritocratic and efficiency-seeking than scandinavians? J. Political Econ. 128, 1753–1788.
Alsan, M., Braghieri, L., Eichmeyer, S., Kim, J., Stantcheva, S., Yang, D., 2023. Civil liberties in times of crisis. AEJ: Appl. Econ. 15 (4), 389–421.
Ambuehl, S., Bernheim, B.D., 2021. Interpreting the will of the people: A positive analysis of ordinal preference aggregation. NBER Working Paper 29389.
Ambuehl, S., Bernheim, B.D., Ockenfels, A., 2021. What motivates paternalism? An experimental study. Am. Econ. Rev. 111 (3), 787–830.
Arad, A., Rubinstein, A., 2018. The People's perspective on libertarian-paternalistic policies. J. Law Econ. 61 (2), 311–333.
Arlegi, R., Bourgeois-Gironde, S., Hualde, M., 2022. Attitudes toward choice with incomplete preferences: An experimental study. J. Econ. Behav. Organ. 204, 663–679.
Arlegi, R., Nieto, J., 2001a. Incomplete preferences and the preference for flexibility. Math. Social Sci. 41 (2), 151–165.
Arlegi, R., Nieto, J., 2001b. Ranking opportunity sets: An approach based on the preference for flexibility. Soc. Choice Welf. 18 (1), 23–36.
Arrow, K.J., 1995. A note on freedom and flexibility. In: Basu, K., Pattanaik, P.K., Suzumura, K. (Eds.), Choice, Welfare, and Development, a Festschrift in Honour of Amartya K. Sen. Oxford University Press, Oxford.
Barberà, S., Bossert, W., Pattanaik, P.K., 2004. Ranking sets of objects. In: Barberà, S., Hammond, P., Seidl, C. (Eds.), Handbook of Utility Theory. Springer, pp. 893–977.
Bartling, B., Cappelen, A.W., Hermes, H., Tungodden, B., 2023. Free to fail? Paternalistic preferences in the United States. Discussion Paper SAM 09/2023, NHH Dept. of Economics.
Bartling, B., Fehr, E., Herz, H., 2014. The intrinsic value of decision rights. Econometrica 82 (6), 2005–2039.
Bauer, M., Chytilová, J., Miguel, E., 2020. Using survey questions to measure preferences: Lessons from an experimental validation in Kenya. Eur. Econ. Rev. 127, 103493.
Baujard, A., 2007. Conceptions of freedom and ranking opportunity sets – a typology. Homo Oeconomicus 24, 231–254.
Benjamin, D.J., Heffetz, O., Kimball, M.S., Szembrot, N., 2014. Beyond happiness and satisfaction: Toward well-being indices based on stated preference. Am. Econ. Rev. 104 (9), 2698–2735.
Bervoets, S., Gravel, N., 2007. Appraising diversity with an ordinal notion of similarity: an axiomatic approach. Math. Social Sci. 53 (3), 259–273.
Blanco, M., Engelmann, D., Koch, A.K., Normann, H.-T., 2014. Preferences and beliefs in a sequential social dilemma: a within-subjects analysis. Games Econom. Behav. 87, 122–135.
Bossert, W., Pattanaik, P.K., Xu, Y., 1994. Ranking opportunity sets: an axiomatic approach. J. Econom. Theory 63 (2), 326–345.
Brunswik, E., 1955. Representative design and probabilistic theory in a functional psychology. Psychol Rev 62 (3), 193.
Carter, I., 1999. A Measure of Freedom. Oxford University Press.
Chaturvedi, A., Green, P.E., Caroll, J.D., 2001. K-modes clustering. J. Classification 18, 35–55.
Chernev, A., Böckenholt, U., Goodman, J., 2015. Choice overload: A conceptual review and meta-analysis. J. Consum. Psychol. 25 (2), 333–358.
Costa-Gomes, M., Crawford, V.P., Broseta, B., 2001. Cognition and behavior in normal-form games: An experimental study. Econometrica 69 (5), 1193–1235.
Dowding, K., van Hees, M., 2009. Freedom of choice. In: Anand, P., Pattanaik, P.K., Puppe, C. (Eds.), Handbook of Rational and Social Choice. Oxford University Press, pp. 374–392.
Dreyer, L., Mahler, D., 2022. Free to choose or free to lose? Understanding individual attitudes toward paternalism. Behav. Public Policy 1–23.
Engelmann, D., Strobel, M., 2000. The false consensus effect disappears if representative information and monetary incentives are given. Exp. Econ. 3, 241–260.

Fallucchi, F., Nosenzo, D., Reuben, E., 2020. Measuring preferences for competition with experimentally-validated survey questions. J. Econ. Behav. Organ. 178, 402–423.
Ferreira, J.V., Hanaki, N., Tarroux, B., 2020. On the roots of the intrinsic value of decision rights: Experimental evidence. Games Econom. Behav. 119, 110–122.
Foster, J., 2011. Freedom, opportunity and well-being. In: Arrow, K., Sen, A.K., Suzumura, K. (Eds.), Handbook of Social Choice and Welfare. vol. II, Elsevier, pp. 687–728.
Gaertner, W., Schokkaert, E., 2012. Empirical Social Choice: Questionnaire-Experimental Studies on Distributive Justice. Cambridge University Press, Cambridge.
Gravel, N., 2008. What is freedom? In: Handbook of Economics and Ethics. Edward Edgar Publishing, London.
Gul, F., Pesendorfer, W., 2001. Temptation and self-control. Econometrica 69 (6), 1403–1436.
Haaland, I., Roth, C., Wohlfart, J., 2023. Designing information provision experiments. J. Econ. Lit. 61 (1), 3–40.
Hastie, T., Tibshirani, R., Friedman, J., 2001. The Elements of Statistical Learning. vol. 1, Springer series in statistics, New York.
Hausman, D.M., 2012. Preference, Value, Choice, and Welfare. Cambridge University Press, New York.
Hayek, F.A., 1960. The Constitution of Liberty. Routledge., London.
Hufe, P., Weishaar, D., 2024. Just cheap talk? Investigating fairness preferences in hypothetical scenarios. CRC Discussion Paper 515.
Iriberri, N., Rey-Biel, P., 2013. Elicited beliefs and social information in modified dictator games: What do dictators believe other dictators do? Quant. Econ. 4 (3), 515–547.
Iyengar, S., Lepper, M., 2000. When choice is demotivating: Can one desire too much of a good thing? J. Pers. Soc. Psychol. 79, 995–1006.
Konow, J., 2009. Is fairness in the eye of the beholder? An impartial spectator analysis of justice. Soc. Choice Welf. 33, 101–127.
Kreps, D.M., 1979. A representation theorem for "preference for flexibility". Econometrica 47, 565–577.
Le Lec, F., Tarroux, B., 2020. On attitudes to choice: some experimental evidence on choice aversion. J. Eur. Econ. Assoc. 18 (5), 2108–2134.
Mill, J.S., 1859. On Liberty, 2010th ed. Penguin Classics, London.
Müller, D., Renes, S., 2021. Fairness views and political preferences: evidence from a large and heterogeneous sample. Soc. Choice Welf. 56 (4), 679–711.
Nehring, K., Puppe, C., 2002. A theory of diversity. Econometrica 70 (3), 1155–1198.
Nozick, R., 1974. Anarchy, State, and Utopia. Basic Books.
Pattanaik, P.K., Xu, Y., 1990. On ranking opportunity sets in terms of freedom of choice. Rech. Econ. de Louvain 56, 383–390.
Pattanaik, P.K., Xu, Y., 1998. On preference and freedom. Theory and Decision 44, 173–198.
Pattanaik, P.K., Xu, Y., 2000. On ranking opportunity sets in economic environments. J. Econom. Theory 93 (1), 48–71.
Puppe, C., 1996. An axiomatic approach to "preference for freedom of choice". J. Econom. Theory 68, 174–199.
Rawls, J., 1971. A Theory of Justice. Harvard University Press, Cambridge.
Roemer, J., 1998. Equality of Opportunity. Harvard University Press.
Sen, A.K., 1985. Commodities and Capabilities. North-Holland.
Sen, A.K., 1988. Freedom of choice: Concept and content. Eur. Econ. Rev. 32, 269–294.
Sen, A.K., 1990. Welfare, freedom and choice: A reply. Rech. Econ. de Louvain 56, 451–485.
Sen, A.K., 2002. Rationality and Freedom. Harvard University Press, Cambridge, MA.
Sher, I., 2018. Evaluating allocations of freedom. Econ. J. 128, 65–94.
Stantcheva, S., 2023. How to run surveys: A guide to creating your own identifying variation and revealing the invisible. Annu. Rev. Econ. 15, 205–234.
Sugden, R., 1998. The metric of opportunity. Econ. Philos. 14, 307–337.
Sugden, R., 2004. The opportunity criterion: Consumer sovereignty without the assumption of coherent preferences. Am. Econ. Rev. 94 (4), 1014–1033.
Toussaert, S., 2018. Eliciting temptation and self-control through menu choices: A lab experiment. Econometrica 86 (3), 859–889.
UK Government, 2023. Patients given more choice to help cut NHS waiting times - what you need to know. https://healthmedia.blog.gov.uk/2023/05/25/patients-given-more-choice-to-help-cut-nhs-waiting-times-what-you-need-to-know/.
Van Hees, M., 1998. On the analysis of negative freedom. Theory and Decision 45, 175–197.
World Values Survey, 2022. Inglehart–welzel cultural map. https://www.worldvaluessurvey.org/WVSContents.jsp?CMSID=Findings.