



HAL
open science

Into the Far West? Investigating Health Policy-Makers' Willingness to Adopt Decrementally Cost-Effective Innovations Using a DCE Approach

Ivan Tzintzun, Jonathan Sicsic, Lise Rochaix

► **To cite this version:**

Ivan Tzintzun, Jonathan Sicsic, Lise Rochaix. Into the Far West? Investigating Health Policy-Makers' Willingness to Adopt Decrementally Cost-Effective Innovations Using a DCE Approach. 2023. halshs-04154933

HAL Id: halshs-04154933

<https://shs.hal.science/halshs-04154933>

Preprint submitted on 7 Jul 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



PARIS SCHOOL OF ECONOMICS
ECOLE D'ECONOMIE DE PARIS

WORKING PAPER N° 2023 – 25

**Into the Far West? Investigating Health Policy-Makers' Willingness
to Adopt Decrementally Cost-Effective Innovations Using a DCE
Approach**

Ivan Tzintzun
Jonathan Sicsic
Lise Rochaix

JEL Codes:

Keywords: Discrete Choice Experiment ; Decrementally Cost-Effective Interventions ; Disinvestment, health policy-makers choices.

anr [©]
agence nationale
de la recherche
AU SERVICE DE LA SCIENCE



Into the Far West?

Investigating Health Policy-Makers' Willingness to Adopt Decrementally Cost-Effective Innovations Using a DCE Approach.

Ivan Tzintzun¹, Jonathan Sicsic², and Lise Rochaix^{1,3}

¹Paris School of Economics - Hospinnomics

²Université Paris Cité, LIRAES F-75006

³Paris 1 Pantheon-Sorbonne University

July 6, 2023

Abstract

In a context of increasingly limited resources, a number of strategies, such as the adoption of decrementally cost-effective interventions (d-CEIs), which are both less clinically effective and less costly, could offer potential levers at enhancing both efficiency and equity in healthcare systems. These interventions are located in the South-West (S-W) quadrant of the cost-effectiveness plane, and have yet received little attention from researchers or HTA agencies as they are often perceived as per se “unethical” or “unacceptable”, hence the reference sometimes made to the S-W quadrant as the ‘Far West’.

The purpose of our paper is to investigate policy-makers’ willingness to adopt d-CEIs using a choice experiment. We use a two-stage pairwise DCE survey to elicit (i) preferences for d-CEIs’ attributes in forced choices and (ii) adoption preferences, i.e. the determinants of d-CEIs’ adoption (unforced choices). We investigate the effect (and trade-offs) between three attributes: health loss (very small to significant), reversibility defined as the possibility to switch back to usual care (from possible to hardly possible) and cost-savings (from 5% to 15% of a fixed budget). Such trade-offs are contextualized by using two sensitivity attributes: disease severity (low and moderate) and savings uncertainty (low and high). Our final sample consists of 180 respondents with 46.7% originating from France and the remaining respondents from other EU countries.

All attributes’ levels have a significant effect in the two decision stages. The “health loss” attribute dominates in the first stage followed by “reversibility”: we calculate that decision-makers would require 28.3% increase of budget savings to be indifferent between a scenario of small versus significant health losses and 14.5% budget savings to be indifferent between a scenario of possible and hardly possible reversibility. In contrast, the “reversibility” attribute dominates in the second stage suggesting that anticipated regret may play a role in adoption decisions.

A stratified analysis based on the sensitivity attributes reveals the existence of heterogeneity across respondents' preferences over attributes and determinants of adoption. Our analyses provide original results regarding the inter-individual and practical determinants of policy-makers' willingness to adopt d-CEIs

Keywords: Discrete Choice Experiment, Decrementally Cost-Effective Interventions, Disinvestment, health policy-makers choices.

1

1 Introduction

In healthcare, as in other publicly funded services, policy-makers have to allocate scarce public resources by choosing the most effective interventions, often facing difficult priority-setting dilemmas. Health technology assessment (HTA) bodies support decision-making by providing recommendations on the added value of new health interventions (such as screening or drugs) compared to usual care. Many countries have identified their principles for priority-setting according to the values and principles specified at the EU level. However, practical ways in which these values and principles are embedded in health care systems vary significantly between EU Member States. This is particularly true when deciding which goods and services should be included in the healthcare basket and the mechanisms used for their funding and provision.

In a context of increasingly limited resources, a number of optimization strategies, such as disinvestment, have received growing attention. Disinvestment is defined by Calabrò et al. [2018] as "withdrawing health resources from any existing health care practices, procedures, technologies or pharmaceuticals that are deemed to deliver little or no health gain for their cost, and thus do not represent efficient health resource allocation". The authors analyse the use of disinvestment strategies in a number of EU countries' HTA agencies and provide evidence of this increasing interest. Unlike disinvestment, very little research has been carried out on the more general strategy of adopting decrementally Cost-Effective Interventions, hereafter referred to as d-CEIs. It corresponds to the explicit inclusion, at the assessment onset, of less clinically effective but also less costly interventions, thereby considering d-CEIs as legitimate comparators. They are located in the South-West (S-W) quadrant of the cost-effectiveness plane, and have received much less attention from HTA agencies than incrementally cost-effective interventions (i-CEIs), located in the North-East (N-E) quadrant. In fact, d-CEIs are often perceived as per se "unethical" or "unacceptable", hence the reference being sometimes made to the S-W quadrant as

¹We would like to thank the European Union's Horizon 2020 research and innovation programme for funding this research. We thank the European Health Economics Association (EuHEA) and the valuable discussions by Anthony Lepinteur and Sandy Tubeuf. We also thank the Society for Benefit-Cost Analysis: European Conference 2022 (SBCAEC2022) and our discussant, Elin Vimefall. Finally, we thank the multidisciplinary team that participated in this project, Panos Kanavos, Jean-Claude Dupont, Emelie Lindström and all members from the IMPACT-HTA project.

the 'Far West'.

The potential of d-CEIs and their (explicit) adoption as comparators in economic evaluation is usually ignored by HTA agencies. Indeed, the UK's National Institute for Health and Care Excellence (NICE) guidelines only mention that these interventions should be recommended if they make available sufficient resources that can be re-invested in public sector care or in services to increase the welfare of the population receiving care. Similarly, the EUnetHTA core model recommends that in cases of both lower costs and lower clinical effectiveness, further analysis is necessary before deciding on adoption, and that an intervention located in the S-W quadrant should be considered cost-effective if society considers that the cost savings compensate for the lower effectiveness (EUnetHTA, 2011). However, the core model does not address how this trade-off is to be documented or the strategy implemented.

Yet there is a growing interest in the potential value of d-CEIs at enhancing both efficiency and equity of healthcare systems. A 2020 consultation for reviewing NICE's Centre for Health Technology Evaluation (CHTE) methods for health technology evaluation acknowledged that decisions in the South-West quadrant do happen and that they may indeed provide net health benefits at population level, due to savings' reinvestment to maximise the system's overall health outcomes (NICE [2020]). An example of d-CEI adoption is the recommendation to use converting enzyme inhibitors (ACE inhibitors) as first-line in the treatment of high blood pressure, leaving the more expensive Angiotensin Receptor Blockers (ARB) for those patients who experience a dry cough as side effect. This recommendation, endorsed by a number of HTA agencies (France, Britain, Sweden), is akin to 'stepped care'. The NICE review indicated the need for further guidance for the assessment of this type of intervention to ensure consistency, and also suggested using net benefit approaches, which are considered as more appropriate in case d-CEIs are to be included as comparators.

From an economic point of view, the topic sparks interest since the decision-making process in healthcare resource allocation is complex and may be influenced by various factors, including psychological biases and preferences. Prospect theory, developed by [Kahneman and Tversky \[1979\]](#), offers insights into how individuals make decisions under conditions of uncertainty and risk, which are prevalent in healthcare resource allocation.

According to prospect theory, individuals evaluate outcomes based on their subjective perception of gains and losses, rather than objective measures. This theory suggests that decision-makers tend to be risk-averse when facing potential losses and risk-seeking when facing potential gains. In the context of healthcare resource allocation, policy-makers may exhibit similar biases, preferring interventions that promise significant health gains while being hesitant to adopt interventions that involve potential health losses, even if they are cost-effective.

The purpose of our paper is therefore to understand the obstacles to d-CEIs' adoption by health policy-makers. To the best of our knowledge, no previous study has inquired into health policy-makers' preferences regarding d-CEIs. The

closest studies are by [Rotteveel et al. \[2022a\]](#) and [Rotteveel et al. \[2022b\]](#), but they focus on disinvestment strategies and investigate citizens' preferences in the Netherlands regarding the relative importance of health care interventions' attributes. Our study therefore fills a significant gap in the literature by studying health policy-makers' preferences regarding d-CEIs in a number of countries in Europe in order to identify the conditions that would allow their consideration as comparators in HTAs and their possible adoption in replacement to usual care.

This study contributes to the economics literature by estimating: i) preferences of individuals in a situation where they have to choose between scenarios that entail a health loss; ii) the determinants of willingness to adopt a d-CEI; iii) the compensation requirements for their adoption, in terms of savings, and iv) the comparison of preferences in forced choices (conditional preferences) versus unforced choices (i.e., adoption preferences).

Our DCE yields three main results. First, the main attribute that shapes individual preferences over d-CEIs in the first stage (forced choices) is related to significant health losses. The compensation requirements over d-CEIs preferences suggest that: i) decision-makers are indifferent between a scenario of small and significant health losses only if it is compensated by 28.3% of budget savings; and ii) a d-CEI with possible reversibility and to hardly possible reversibility if it is offset by 14.5 % of budget savings. However, when analyzing the second stage results (i.e., adoption preferences), we find that the attribute "hardly possible to reverse" dominates over the other attributes. Respondents are willing to adopt: i) a d-CEI involving an increased health loss (from very low to significant) only if it is offset by 14.1% budget savings; and ii) moving from immediate reversibility to hardly possible reversibility would require a compensation of 16.5% in budget savings. Second, a stratified analysis of the compensation requirements according to participants' responses of the follow-up questions in the second stage (i.e., "*would you be ready to substitute usual care by the option you selected*" and "*would you be ready to substitute usual care by the option selected if disease severity of the patient -before treatment- changes to moderate*"), reveals the existence of preference heterogeneity across participants.

The remainder of the paper is organized as follows: Section 2 presents the discrete choice experiment. Sections 3 presents the summary statics, the overall description of choices, the frequency of choices, and the descriptive statistics for willingness to adopt. Section 4.1 shows the first stage results, section 4.2 the second stage results, section 4.3 uses the sensitivity attributes to explore preference heterogeneity. Section 5 shows the robustness checks. Finally, section 6 offers a discussion and draws the main conclusions from the study.

2 The discrete choice experiment

We designed a DCE to analyze the main determinants of health policy-makers' willingness to adopt d-CEIs beyond cost-effectiveness. Our DCE presented participants with eight choice sets, representing d-CEIs implying different levels of health losses, decision reversibility and magnitude of savings.

2.1 Experimental design

The DCE design was informed by a number of preparatory steps. A systematic international literature review was carried out to identify the existing level and quality of evidence on d-CEIs and their potential associated savings. A narrative review of existing HTA bodies' guidelines and practices further documented attitudes and obstacles to d-CEIs' adoption. Qualitative interviews were also carried out to inform the overall design of the experiment and the choice of attributes and levels. Potential scenarios were discussed with members of the IMPACT-HTA consortium and the DCE was piloted with master students and participants of a meeting organized with a selected number of health decision-makers in France. Based on findings from the qualitative phase, a number of choices were made regarding the format and design of the DCE.²

First, our DCE follows a two-stage decision-making process. In the first stage ("conditional preference elicitation task"), participants have to choose between two d-CEI alternatives in a given choice task, and therefore they are forced to choose between the two alternatives. In the second stage ('adoption preference task'), participants can choose whether they are ready to replace usual care by the preferred option in the previous stage. This design is also named dual response choice experiment or sequential choice. Examples of this type of experiments are Traets et al. [2022], Soekhai et al. [2019], Cheng et al. [2012], and Kallas and Gil [2012].

We prefer this design over standard pairwise + opt-out task for two main reasons. First, because we expected a limited number of answers (due to the study target, namely healthcare decision makers), and forced choices in the first stage allowed to obtain maximum information about preferences. Indeed, as it is not possible to define the opt-out option based on the attributes' levels we use, non-demanders of d-CEI would not reveal their preferences' weights for the attributes. Second, we aimed to differentiate valuation of conditional preferences (first stage elicitation tasks) and adoption preferences (second stage elicitation tasks). In particular, we expected that decision-makers could be more risk averse in adoption preferences due to the change in reference point where the "current situation" refers to higher health gains. The two (forced and unforced) elicitation tasks were repeated eight times to offer different combinations of attributes and levels (see figure 5 and 6). In detail, sixteen optimal combinations of attributes and levels were selected, based on an efficient fractional design and allocated in eight pairwise choice tasks, using NGENE software (Choice metrics). Because of the difficulty to obtain prior estimates from a pilot study, we defined theoretical priors consistent with prospect theory (i.e., loss aversion, reference dependence). We hypothesized that increasing health loss levels and reversibility delays would have negative effects on adoption choices while increasing budget savings would increase utility but at lower rate (due to loss aversion).

Second, in order to maximize consequentially, we chose a role-play design,

²The report delivered to the IMPACT-HTA consortium is available online and can be accessed here: https://www.impact-hta.eu/_files/ugd/e1a359_0ebaed9d87624755b875eb2cbe8571da.pdf?index=true

with respondents being asked to act as a regional health policy officer in charge of a set budget of €100 million per year and in charge of a population of 10,000 patients suffering from a given (unspecified) health condition.

Third, the DCE was kept generic, with no reference to specific interventions or illnesses, in order to avoid strong anchoring effects, as recommended in the experimental economics literature.

Fourth, we address the potential bias introduced by respondents who do not want to participate by offering an ‘escape’ route at an early stage of the DCE to ensure that these respondents would at least answer a minimal set of questions. This allowed distinguishing between respondents objecting to the duration of the experiment, to the experimental design itself or to the actual focus of the study, i.e. d-CEIs. Participants who drop out due to the latter will minimize the “no-protest” bias.

Fifth, the qualitative analysis and pilot tests performed prior to the DCE yielded another important finding, i.e. that participants would likely refuse to ‘play the game’ for high levels of disease severity or high uncertainty associated with the cost savings. To improve DCE uptake and response quality, we chose to study preferences over a limited range of the SW quadrant. Participants were therefore systematically reminded at the beginning of each choice set that patients only experienced low disease severity and that uncertainty related to costs savings was also low.

The DCE was implemented online, using Qualtrics software, and targeted EU respondents from HTA agencies or health technologies’ pricing and reimbursement committees, Officers of governmental or non-governmental organisations, representatives of citizens or consumer-patient advocacy groups, health economists (both senior and in training).

2.2 Attributes and levels

A large number of attributes were identified in the preparatory stage and choices were made to narrow them down to a manageable size, in terms of length of time for respondents. Three semi-structured interviews were conducted with health policy-makers to specifically inform the choice of attributes and levels and a qualitative analysis was carried out on material collected during DCE completion in seven individual think-aloud sessions. The experiment focused on those attributes which would be most relevant to health policy-makers in their quest to document society’s willingness to consider the clinical losses as equivalent to the cost savings, as suggested in the EUNETHTA guidance.

The three attributes were defined as follows:

1. **Health loss:** Health loss refers to a loss in quality of life as perceived by the patient population. This could take the form of side effects (such as coughing and/or additional constraints associated with taking the treat-

ment (such as a pill every day rather than every week). The levels considered are: Very small, small and significant.

2. **Possibility to switch back (reversibility):** Possibility to switch back from decremental cost-effective intervention to usual care (expressed in time). The levels considered are: possible at any time; possible after some delay; hardly possible after long delays.
3. **Cost-savings:** Annual savings being made when switching from usual care to a decremental cost-effective intervention. The savings are defined as a percentage of a fixed healthcare budget of €100 million per year. The levels considered are: 5%, 10%, 15%.

In order to investigate the implications of having set severity (respectively uncertainty about savings) as low in the DCE, two secondary attributes were analysed through two follow-up questions asked after choice task 2 and 7. We ask respondents whether they would reconsider their decision of adopting the d-CEI in case (i) disease severity and (ii) uncertainty about the cost-savings were to increase (see figure 7). While these additional questions do not yield preference weights, compared to the three chosen attributes, this sensitivity attributes allows capturing the potential impact of increasing severity (respectively uncertainty) on preferences and health decision-makers' willingness to adopt their preferred option.

The secondary attributes for sensitivity analysis are defined as follows:

1. **Disease severity:** The disease severity attribute relates to the patient experience of the disease, as indicated by their EQ-5D Quality of Life Score. The levels considered are: low and moderate.
2. **Uncertainty:** The uncertainty attribute is based on the probability of observing the expected cost-savings if adopting the decremental cost-effective intervention. The levels considered are: low and high. Respondents were informed that uncertainty indicates the possibility that the d-CEI will not bring the expected cost-saving.

2.3 Estimation method

The choice data were analyzed within a random utility maximization framework i.e., assuming a random utility function for each hypothetical scenario (including a systematic and unobserved portion of utility) and assuming utility maximization decision rule (McFadden et al. [1973]). We used a mixed multinomial logit (MIXL) model specification, allowing the preference parameters to be randomly distributed across the sample and thus accounting for (i) unobserved preference heterogeneity and (ii) correlation of choices between participants (McFadden and Train [2000]). For respondent n , alternative j , and choice situation t , utility U , can be given by

$$U_{n,j,t} = V_{n,j,t} + \epsilon_{n,j,t} \quad (1)$$

where, $U_{n,j,t}$, includes a deterministic component $V_{n,j,t}$, and a random component $\epsilon_{n,j,t}$, which is assumed to be an independently and identically distributed Extreme Value Type I function (Hensher et al. [2005]; Manski [2001]),

thus leading to the multinomial logit choice specification (McFadden et al. [1973]). Further, the deterministic part of the utility can be re-written as:

$$V_{n,j,t} = f(\beta_n, x_{n,j,t}) \quad (2)$$

where β_n is the vector of sensitivities for the respondent, and $x_{n,j,t}$ is a vector of attributes for alternative j . In this DCE application, the deterministic utility is given by:

$$\begin{aligned} V_{n,j,t} = & \beta_{1,n}HealthLossSmall_j + \beta_{2,n}HealthLossSignificant_j \\ & + \beta_{3,n}ReversibilityWithDelay_j + \beta_{4,n}ReversibilityHardlyPossible_j \\ & + \beta_{5,n}Savings_j \end{aligned} \quad (3)$$

where $U_{n,j,t}$ is the utility a decision maker n derives from choosing alternative (scenario) j in choice task t ; $HealthLossSmall_j$, $HealthLossSignificant_j$, $ReversibilityWithDelay_j$, $ReversibilityHardlyPossible_j$ represent dummy coded attributes' levels displayed in scenario j (level 1 of all attributes was used as reference), $Savings_j$ represent the level of savings in scenario j , $\beta_{1,n}, \dots, \beta_{5,n}$ are the respective random effects (part-worth utilities) associated with each attribute level, the subscript n denoting respondent-specific parameters.

We estimate the first stage parameters by using a mixed logit model (MIXL), using 500 Halton draws with correlated random coefficients between all 5 attributes' levels, assuming normal distribution for each parameter $\beta_{k,n}$. Though particularly computationally intensive, this model is known to be most flexible (Hess and Train [2017]). In particular, it allows accounting for scale heterogeneity, i.e., various degrees of consistency of decisions across respondents (Hess and Rose [2012]).

In order to visually represent the degree of preference heterogeneity, we predicted individual-level coefficients from the MIXL model (i.e. $\hat{\beta}_{k,n}, k = 1, \dots, 5$) using the methodology detailed in Revelt and Train [1998].

We estimate the second stage parameter, e.g. the impact of each attribute on the willingness to substitute usual care, using a random effects linear probability model, with the response to the question "would you be willing to substitute usual care by the option you selected?" as dependent variable, and the attributes of the selected option as independent (exogenous) variables. Note that because of the experimental design (zero correlation between attributes and individual characteristics), the results of these models are not significantly different from a fixed effects specification.

3 Data

Respondents were recruited from January to April 2021, based on a number of recommendations from participants to the EU IMPACT-HTA project. A total of 234 questionnaires were completed. Of these, 54 were excluded from the analysis in line with the exclusion criterion described above. The overall

characteristics of the 180 respondents are summarized in table 1.

46.7% of respondents were French, the remaining participants originating from various EU countries: the second largest group being from Poland (8.3%) and the third largest from the UK (7.8%). The sample is balanced across genders. Regarding age, the largest group has participants between 18 and 35 (43.7%), followed by individuals age 36 to 55 (%). Senior health economists (27%) and students (24.2%) represented the two largest groups, with half of the sample having an economics background (52.2%).

Question	Percentage	N
Country of residence		180
France	46.70%	
Poland	8.30%	
Slovenia	6.10%	
Switzerland	5.6%	
UK	7.80%	
Other countries	25.60%	
Gender		178
Female	52.80%	
Male	42.70%	
Non-binary	0.56%	
Prefer not to say	3.93%	
Age		174
18 to 35	43.70%	
36 to 55	43.10%	
56 to 65	10.30%	
Older than 65	2.90%	
Current position		178
Health technology assessment professional	18.50%	
Member of health technologies; pricing and reimbursement committee	3.40%	
Officer of government and non-government organizations	6.20%	
Representative of citizens or consumer / Patient advocacy group	1.10%	
Health Economist	27.00%	
Student / in training	24.20%	
Clinician	7.30%	
Other	12.40%	
Studies (can be cumulated)		180
Medical Studies	21.70%	
Economics Studies	52.20%	
Public Health Studies	23.30%	
Pharmaceutical Studies	5.00%	
Public Policy Studies	12.80%	
The experiment was technically easy to follow		178
Strongly disagree	1.69%	
Disagree	5.06%	
Neither agree nor disagree	5.06%	
Agree	49.44%	
Strongly agree	38.76%	

Table 1: Summary statistics I - Sociodemographic characteristics

Summary statistics in table 2 show that around 61.67% of respondents al-

ready knew about d-CEIs but most participants did not know whether their country had ever adopted a d-CEI (36.94%). Additionally, 75.68% of the sample had no direct experience of adopting cost-effective interventions. Interestingly, from a patient perspective, respondents declared willingness to adopt in 45% of cases in case d-CEI adoption were prescribed by their doctor. Statistics show that most participants did not change their views on d-CEIs after the experiment (38.76%). Finally, most participants had a positive view on d-CEIs: 63.48% thought that such interventions may contribute to a better allocation of resources; 52.55% considered that HTA committees should consider these interventions as comparators.

Question	Percentage	N
Did you know about d-CEIs before watching the introduction video?		180
Yes	61.67%	
No	38.33%	
Does your country adopt d-CEIs ?		111
Yes, regularly	4.50%	
Yes, sometimes	36.04%	
No	22.52%	
I don't know	36.94%	
Do you have experience with adopting d-CEIs ?		111
A lot of experience	2.70%	
Some experience	18.02%	
No experience	75.68%	
I don't know	3.60%	
This experiment changed my opinion on d-CEIs		178
Strongly disagree	10.11%	
Disagree	24.72%	
Neither agree nor disagree	38.76%	
Agree	18.54%	
Strongly agree	7.87%	
As a patient, I may accept a d-CEIs option if it was offered to me by my doctor as a suitable replacement for usual care:		178
Strongly disagree	8.43%	
Disagree	11.80%	
Neither agree nor disagree	22.47%	
Agree	45.51%	
Strongly agree	11.80%	
Health care resources should be allocated with the objective of maximizing overall health of the population		178
Strongly disagree	1.12%	
Disagree	1.69%	
Neither agree nor disagree	8.43%	
Agree	48.88%	
Strongly agree	39.89%	
d-CEIs may contribute to a better allocation of the health care budget		178
Strongly disagree	2.25%	
Disagree	3.37%	
Neither agree nor disagree	10.67%	
Agree	63.48%	
Strongly agree	20.22%	
It is acceptable to consider d-CEIs in HTA Committees		178
Strongly disagree	1.12%	
Disagree	6.74%	
Neither agree nor disagree	10.67%	
Agree	52.25%	
Strongly agree	29.21%	
Acting as a local health officer, would you say that the COVID-19 crisis has made you....		176
...more ready to adopt d-CEIs because money has become more scarce	43.18%	
...less ready to adopt d-CEIs as the crisis has shown that money can always be found when needed	23.30%	
Neither	33.52%	

Table 2: Summary statistics II - Participants attitudes toward d-CEIs

3.1 Description of the choice tasks and frequency of choices

Table 3 shows the eight different choice tasks and the distribution of choices between scenario 1 and scenario 2. As a rationality test, we included choice tasks 4 and 8, in which scenario 2 dominates over scenario 1 and thus we expected participants to chose accordingly. We observed that 7% failed the rationality test in choice task number 4 (3% of our sample failed the test). In total, this is equivalent to 14 out of 180 individuals in our sample (7.78%).

Choice task	Scenario 1			Scenario 2			Choice statistics	
	Health loss	Reversibility	Savings	Health loss	Reversibility	Savings	choices (N)	Scenario 1 chosen*
1	Small	Possible, anytime	10%	Very small	Hardly possible	10%	199	74%
2	Very small	Hardly possible	5%	Significant	Possible, anytime	15%	182	59%
3	Significant	Possible, anytime	5%	Very small	Possible, delays	15%	182	5%
4	Small	Possible, delays	10%	Very small	Possible, anytime	10%	185	7%
5	Significant	Possible, delays	15%	Small	Possible, anytime	5%	181	15%
6	Small	Hardly possible	15%	Significant	Possible, anytime	5%	184	84%
7	Very small	Possible, delays	5%	Small	Hardly possible	10%	182	66%
8	Significant	Hardly possible	10%	Small	Possible, delays	15%	152	3%

Table 3: Description of choice tasks and choice frequency. Participants had to chose between two d-CEIs scenarios in each choice task. Scenario 2 dominates in choice task number 4 and 8 and thus represents a rationality test.

Based on the statistics in table 3, we find the following. First, when presented with a choice task that compares a significant health loss to a small or very small one (i.e. choice task 2, 3, 5, 6 and 8), we observe on average that only 10.3% of respondents will prefer the scenario with significant health losses. Second, when participants face a choice task that compares hardly possible reversibility with possible at any time or with delays (i.e. choice task 1,2,6,7, and 8), results in table 3 suggest that around 26% of respondents will prefer the scenario with hardly possible reversibility. And third, when facing a choice task that compares 5% savings with 10% or 15%, around 33% of participants will chose the scenario with the lowest savings. Therefore, the summary statistics suggest that participants are more adverse to significant health losses, followed by hardly possible reversibility and 5% savings.

3.2 Descriptive statistics for willingness to adopt

Table 4 shows the descriptive statistics for the follow-up question that enquires into the willingness to adopt (i.e. substitute usual care with a d-CEI). For instance, we observe that 85% of participants were ready to substitute usual care with the selected option in the default scenario (low disease severity and low uncertainty). Overall, there was a relatively high willingness to substitute usual care with the selected hypothetical option, from a minimum of 60% in task 2 to

a maximum of 94 % in task 4 (Table 4).

Based on the statistics of the follow-up questions in choice task 2 and 7, we observe that the decision to set disease severity and uncertainty related to cost-savings as low are important in ensuring willingness to adopt d-CEIs. We observe two main result in table 4. First, when disease severity of the patient changed from low to moderate: the willingness to substitute decreases from 60% to 40% in task 2, and from 79% to 50% in choice task 7. Second, when uncertainty related to cost-savings changed from low to high, the willingness to substitute decreases from 60% to 22% in task 2, and from 79% to 26% in task 7. Differences in the willingness to substitute if disease severity changed from low to moderate in choice task 2 and 7 may be related to the fact that the health losses attributes in choice task 2 give two possible choices to participants: "very small" vs "significant"; on the contrary, choice task 7 considers "very small" vs "small" health losses. Since the point of reference changed (e.g. disease severity), it is intuitive that individual will be more likely to substitute under choice task 7.

Choice task	Willigness to substitute usual with selected option if:		
	Low disease severity and low uncertainty (%)	Disease severity of the patient changes to moderate (%)	Uncertainty related to cost-savings changes to high (%)
1	85%	-	-
2	60%	40%	22%
3	89%	-	-
4	94%	-	-
5	76%	-	-
6	61%	-	-
7	79%	50%	26%
8	79%	-	-

Table 4: Descriptive statistics on the follow-up question (sensitivity attributes): willingness to substitute with usual care. N=180

By using the follow-up questions in choice task number 2, table 13 in annex D shows personal characteristics mean values of four different groups of participants: 1) respondents who are willing to adopt the d-CEI if disease severity increases to moderate; 2) respondents who are not willing to adopt the d-CEI if disease severity increases to moderate; 3) respondents who are willing to adopt the d-CEI if savings uncertainty increases to moderate; 4) respondents who are not willing to adopt the d-CEI if savings uncertainty does not increases to moderate. Overall, the results suggest there are no significant differences between groups. Nonetheless we find two exceptions. First, compared to respondents that are willing to adopt d-CEI if savings uncertainty increases to high, respondents who are not willing to adopt are more likely to be French (15%). Second, respondents who are willing to adopt when disease severity increases to moderate to high are more likely to adopt if savings uncertainty increases to high and vice-versa.

3.3 Hypotheses

The main question addressed in this study is under which conditions will a health policy-maker consider, and possibly adopt a d-CEI. It bridges over to the

more general issue of the relationship between individual and collective benefits: under which conditions should the collective interest override the individual interest?

The following research hypotheses guided the analysis of the data:

1. Based on the summary statistics in section 3, we expect respondents to prioritize either the attribute on the health loss (as perceived by the patient population) or the possibility to switch back to usual care (expressed in time), and cost-savings will be considered secondary (least important of the attributes).
2. Respondents will be more willing to adopt interventions if the health loss (as perceived by the patient population) is low and if the possibility to switch back to usual care (expressed in time) is possible, at any time.
3. Respondents will be more likely to consider adoption if the two sensitivity attributes, disease severity and uncertainty of expected savings) are low, with more importance placed on the former.
4. Compared to respondents who are willing to adopt their preferred d-CEI when the sensitivity attributes increase, respondents who are not willing to adopt will be more likely to require a higher compensation in terms of savings in order to forego health or possibility to reverse.

4 Results

4.1 First stage estimates: conditional preference elicitation task

4.1.1 Results on main attributes levels

This section presents the part-worth utilities, e.g. $\beta_{i,n}$ in equation 3 related to forced choices (i.e. pairwise comparisons). In the mixed logit models with correlated random coefficients, four attribute levels had negative and statistically significant part-worth utilities, in line with a priori expectations (Table 5): small health losses ($\beta=-2.116$), significant health losses ($\beta=-7.845$), reversibility possible with delays ($\beta=-2.207$), reversibility hardly possible due to long delays ($\beta=-4.027$). In other words, all other measures being equal, respondents' likelihood to choose one scenario over the other decreased with these attribute levels (compared to the reference). On the contrary, the level of cost savings had a positive and significant impact on choices: every 5% increase in savings from the fixed budget increased utility by 1.385.

Dependent variable: choice between scenario 1 or 2	Mean		SD	
	Est	Std. Err	Est	Std Err.
Health loss: very small	<i>Ref</i>		<i>Ref</i>	
Health loss: small	-2.116***	(0.432)	1.255***	(0.377)
Health loss: significant	-7.845***	(1.039)	2.701***	(0.816)
Reversibility: possible, at any time	<i>Ref</i>		<i>Ref</i>	
Reversibility: possible, with delay	-2.207***	(0.582)	1.084**	(0.472)
Reversibility: hardly possible	-4.027***	(0.561)	0.906**	-0.446
Savings (linear, per 5% of budget)	1.385***	(0.269)	-0.902***	(0.208)
N (individuals)	180			
Choice observations	1,408			
Log-likelihood	-518.420			

Standard errors in parenthesis
 *** p<0.01, ** p<0.05, * p<0.1.

Table 5: First stage results - MIXL model with correlated random coefficients with 500 Halton draws.

Significant preference heterogeneity was found for all attribute levels, with statistically significant standard deviations of the underlying normal random effects. The graphical distribution of the predicted individual level coefficients (β) confirms significant heterogeneity in preferences for the above-mentioned attribute levels (see Figure 1).

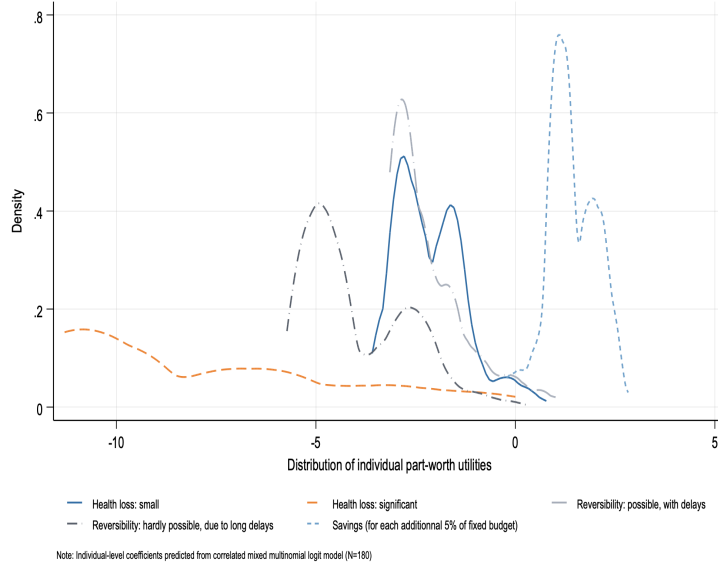


Figure 1: Distribution of individual-level coefficients derived from the first stage correlated MIXL model

The overall importance of attributes is displayed in figure 2. The attribute "Health loss: significant" had the greatest influence (normalized logworth = 100%, reference), followed by "reversibility: hardly possible" (89%), "Savings"

(74%), "health loss: small" (71%), and "reversibility: possible with delays" (66%).

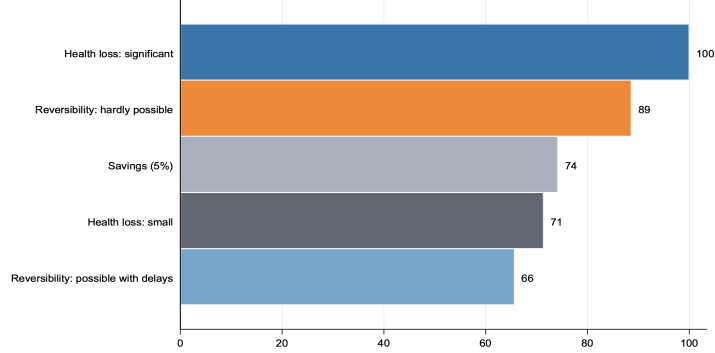


Figure 2: First stage normalized logthworth (%) - Importance of the attributes

4.1.2 Compensation requirements

An important determinant of health policy makers' decision to adopt d-CEIs is the level of health loss and the degree of reversibility of the decision with respect to savings. We further investigate trade-offs between savings and these two other attributes (health losses and decision reversibility) by calculating the marginal rate of substitution as follows:

$$MRS_{i,j} = \frac{\partial U}{\partial X_i} / \frac{\partial U}{\partial X_j} = \frac{\beta_i}{\beta_j} \quad (4)$$

Equation 4 calculates the marginal rate of substitution between two attributes X_i, X_j , where U is the utility function defined in equation 3. We focus here on the additional savings that health decision-makers should be compensated with in order to remain on the same utility level when facing either health losses or reversibility delays. We calculate point estimates based on the β_i parameters in table 5: "health loss: small" (β_1); "health loss: significant" (β_2); "reversibility: with delays" (β_3); and "reversibility: hardly possible" (β_4). In order to be interpreted as the effect of a 1% budget saving (which is equivalent to directly coding the budget saving attributes in unit percentage increases), the savings coefficient (β_j) is divided by 5.

Results in table 6 show that health policy-makers are indifferent between: i) a scenario of "small" and "very small" health losses if it is compensated by a 7,6% of fixed budget savings; ii) a scenario of "significant" and "very small" health losses if it is compensated by a 28,3% of fixed budget savings; iii) a scenario of "reversibility: possible with delays" and "possible at any time" only if it is compensated by an 8% of fixed budget savings; and iv) a scenario of "hardly possible due to long delays" and "possible at any time" if it is compensated by a 14,5% of fixed budget savings.

Attributes and levels	Compensation in terms of savings (%)
Health loss: low	7.6 [5.4, 9.9]
Health loss: significant	28.3 [21.2,35.4]
Reversibility: possible with delay	8.0 [3.3,12.7]
Reversibility: hardly possible	14.5 [10.2,18.9]

Table 6: First stage compensation requirements in terms of savings. 95% confident intervals in brackets (N=180)

4.2 Second stage estimates: adoption preference task

4.2.1 Results on main attributes levels

Table 7 presents the marginal effects of attributes (and their levels) on the probability to “be willing to substitute” (i.e. adoption preferences). Results in table 7 show the random and fixed effect model estimates. Note that because of the experimental design (exogeneity), the results of these models are not significantly different from a fixed effects specification. Results show that the willingness to substitute usual care varies according to the attributes of the d-CEIs. The random effects model (column 1) indicates that the willingness to substitute decreases by 11.9 percentage points (pp) if the health loss was small; 25.9 pp if the health loss was significant (26.5 pp in the FE model), 9.1 pp if reversibility was possible but with delays, 30.3 pp if reversibility was hardly possible due to long delays (30.8 pp in the FE model), and increases 9.2 pp for each additional 5% increase in the fixed budget savings (8.6 pp in the FE model).

	(1)	(2)
	RE	FE
Health loss: very small	<i>Ref</i>	<i>Ref</i>
Health loss: small	-0.119*** (0.022)	-0.119*** (0.022)
Health loss: significant	-0.259*** (0.038)	-0.265*** (0.039)
Reversibility: always possible	<i>Ref</i>	<i>Ref</i>
Reversibility: with delay	-0.091*** (0.024)	-0.091*** (0.024)
Reversibility: hardly possible	-0.303*** (0.025)	-0.308*** (0.025)
Savings (linear 5% of budget)	0.092*** (0.014)	0.086*** (0.014)
Constant	0.787*** (0.030)	0.800*** (0.026)
Observations	1408	1408
N(Individuals)	180	180
Individual Fixed Effect	No	Yes

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 7: Second stage results - Random and fixed effect model estimates of the impact of attributes on willingness to substitute preferred d-CEI for usual care.

The overall importance of attributes is displayed in Figure 3. In contrast to Figure 2, the attribute "reversibility: hardly possible" had the greatest influence (normalized logworth = 100%, reference), "health loss: significant" (80%), "Savings" (79%), "health loss: small" (75%), and "reversibility: possible with delays" (64%). These results imply that the importance of attributes for the part-worth utilities do not necessarily translate into willingness to substitute for standard care.

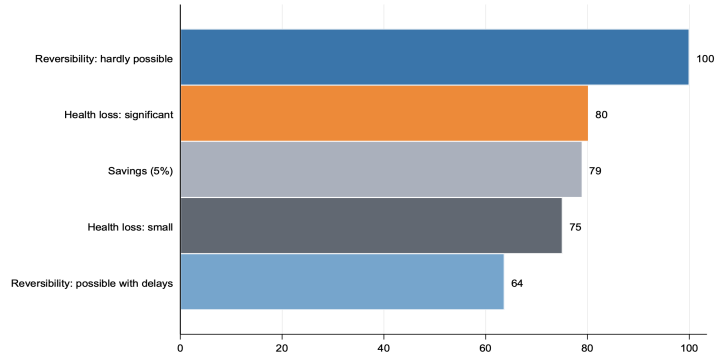


Figure 3: Second stage normalized logthworth (%). Importance of the attributes for willingness to substitute standard care for d-CEI.

4.2.2 Compensation requirements

Results in Table 8 show that, given the initial situation, health policy-makers are willing to substitute for a d-CEI with “small” health losses if this decision is compensated by a 6,5% of fixed budget savings; ii) a scenario of “significant” and “very small” health losses if it is compensated by a 14,1% of fixed budget savings; iii) a scenario of “possible reversibility with delays” and “possible at any time” if it is compensated by a 5.05% of fixed budget savings; and iv) a scenario of “hardly possible due to long delays” and “possible at any time” if it is compensated by a 16,5% of fixed budget savings.

Attribute and levels	Compensation in terms of savings (%)
Health loss: low	6.5 [4.2,8.7]
Health loss: significant	14.1 [9.6,18.6]
Reversibility: possible with delay	5.05 [2.6,7.4]
Reversibility: hardly possible	16.5 [11.7,21.2]

95% confidence intervals in brackets

Table 8: Second stage compensation requirements in terms of savings. 95% confident intervals in brackets (N=180).

4.3 Sensitivity attributes: changes in the reference point

This section uses the sensitivity attributes to study the impact of changes in the reference point. We conduct a stratified analysis of the compensation requirement estimates from the first and second stage of the DCE ³

³Due to its economic relevance we only show the compensation requirements estimates and we do not include the mixed logit results and the random effects models with the impact of attributes on forced preferences (i.e. first stage) and on willingness to adopt (i.e. second stage). The results are available upon request.

The stratification is carried out by classifying participants according to their responses to the follow-up questions in the second stage of choice task 2⁴ in table 3 : a) "Would you be ready to substitute usual care by the option you selected if disease severity of the patient (before treatment) changes to moderate"; and b) "Would you be ready to substitute usual care by the option selected if disease severity of the patient (before treatment) changes to moderate".

These results are relevant because the stratification captures potential behavioral changes between individuals who may respond differently when we modify the design of the DCE by changing the point of reference.

4.3.1 First stage compensation requirements' stratified analysis

Table 9 shows the compensation requirements estimates for the different sub-populations when individuals are forced to choose a d-CEI. Column 1 presents the full sample results discussed in table 6; columns 2 and 3 show the stratified analysis, conditional on whether participants are willing to substitute when disease severity changes from low to moderate; and columns 4 and 5 present the results from the stratified analysis, conditional on whether participants are willing to substitute when savings' uncertainty changes from low to high.

The results suggest some degree of heterogeneity across respondents, conditional on their willingness to substitute. Coefficients in column 2 are always smaller than coefficients in column 3. We find similar results when we compare coefficients in columns 4 and 5. This means that: i) individuals who are not willing to substitute when disease severity increases to moderate require a higher compensation in savings to remain indifferent, for a given change in an attribute level; and ii) individuals who are not willing to substitute when savings' uncertainty increases to high require a higher compensation in savings to remain indifferent, for a given change in an attribute level.

⁴We conducted the same analysis using the follow-up questions in choice task 7. The results offer similar insights and are not included but are available upon request.

	Full sample	Moderate disease severity		High savings uncertainty	
	(1)	(2)	(3)	(4)	5)
Health loss: low	7.6 [5.4,9.9]	6.5 [4,9.1]	9.1 [6.4,11.8]	9 [5.5,12.5]	6.7 [3.8,9.5]
Health loss: significant	28.3 [21.2,35.4]	21.8 [15.6,28]	28.8 [21.3,36.4]	20.9 [14.9,26.9]	36.9 [20,53.9]
Reversibility: possible with delay	8 [3.3,12.7]	4.8 [-0.6,10.2]	8.7 [3.5,14]	4.3 [-1.9,10.6]	9.8 [2.1,17.5]
Reversibility: hardly possible	14.5 [10.2,18.9]	9.3 [5.2,13.3]	14.9 [10.3,19.6]	11.8 [7.1,16.6]	18.1 [9.3,27]
N(Individuals)	180	73	107	40	140
Conditional willingness to substitute		Yes	No	Yes	No

Table 9: First stage compensation requirements in terms of savings stratified according to willingness to substitute when disease severity increases to high. Column 1 is the benchmark result and presents the compensation requirements using the full sample. Columns 2 corresponds to those who are willing to substitute if disease severity changes from low to moderate, while column 3 corresponds to those who are not. Similarly, column 4 corresponds to those who are willing to substitute if savings' uncertainty changes from low to high, while column 5 corresponds to those who are not. 95% confident intervals in brackets.

Table 11 in annex C.1 shows the difference between coefficients in column 2 and 3, and attributes in column 4 and 5 and provides the p-values for the hypothesis that coefficients are indeed different. The results partially confirm the existence of heterogeneity between groups. In particular, we observe that: i) compared to individuals who are willing to substitute when disease severity increases to moderate, individuals who are not willing to substitute require additional 7.0 % in terms of savings to remain indifferent when the health loss attribute changes from very low to significant, and 16.0 % for those that are not willing to substitute when savings' uncertainty increases to high; and ii) compared to individuals who are willing to substitute when savings' uncertainty increases to high, individuals who are not willing to substitute require additional 5.6% in terms of savings to remain indifferent when the reversibility attribute changes from "always possible" to "hardly possible".

4.3.2 Second stage's compensation requirements stratified analysis

Table 10 shows the compensation requirements estimates for the different sub-populations based on the second stage. Column 1 are the full sample results discussed in table 8; columns 2 and 3 shows the stratified analysis conditional on whether participants are willing to substitute when disease severity changes from low to moderate; and columns 4 and 5 stratified the analysis conditional on whether participants are willing to substitute when savings uncertainty changes from low to high.

Our findings indicate some heterogeneity across respondents conditional on their willingness to substitute. In contrast to a situation when individuals are forced to choose, coefficients in column 2 are not always smaller than coefficients in column 3. However, we find similar results when we compare coefficients in column 4 and 5. Despite having similar insight to the first stage, a notorious

difference we find with the second is the relative size of the difference between the coefficients when health loss increases from very low to significant: coefficient in column 3 is more than twice the size of the coefficient in column 2; coefficient in column 5 is almost 9 times the size of coefficient in column 4.

	Full sample	Moderate disease severity		High savings uncertainty	
	(1)	(2)	(3)	(4)	(5)
Health loss: low	6.5 [4.2,8.7]	8.2 [-1.7,18]	6.3 [4.1,8.5]	3.3 [-2.8,9.3]	7.0 [4.6,9.4]
Health loss: significant	14.1 [9.6,18.6]	7.3 [-6.6,21.2]	16.4 [11.5,21.3]	2.1 [-7,11.2]	18.0 [12.4,23.7]
Reversibility: possible with delay	5.05 [2.6,7.4]	13.4 [-2.1,28.9]	3.9 [1.6,6.2]	8.7 [0.6,16.8]	4.4 [1.9,6.9]
Reversibility: hardly possible	16.5 [11.7,21.2]	24.4 [-2.9,51.7]	15.9 [11.4,20.3]	15.8 [2.8,28.8]	16.5 [11.5,21.4]
N(Individuals)	180	73	107	40	140
Conditional willingness to substitute		Yes	No	Yes	No

Table 10: First stage compensation requirements in terms of savings stratified according to willingness to substitute when disease severity increases to high. Column 1 is the benchmark result and presents the compensation requirements using the full sample. Columns 2 corresponds to those who are willing to substitute if disease severity changes from low to moderate, while column 3 corresponds to those who are not. Similarly, column 4 corresponds to those who are willing to substitute if savings' uncertainty changes from low to high, while column 5 corresponds to those who are not. 95% confident intervals in brackets.

Table 12 in annex C.2 shows the difference between coefficients in column 2 and 3, and attributes in column 4 and 5 and provides the p-values for the hypothesis that coefficients are indeed different. The results confirms the existence of heterogeneity between groups only for the "health loss: significant" attribute: i) compared to individuals who are willing to substitute when disease severity increases to moderate, individuals who are not willing to substitute require additional 9.1 % in terms of savings to remain indifferent when the health loss attribute changes from very low to significant, and 15.9 % for those that are not. We cannot reject the hypothesis that there are not significant differences between the rest of attributes.

5 Robustness checks

5.1 Additional tests: non-linearities

5.1.1 First stage: mixed logit estimates

We tested the robustness of our results to deviations from the linearity assumption for the cost-savings attribute, but the results did not suggest the need for more complex (e.g., categorical or polynomial) specifications in terms of model performance (see Table 14). We performed two additional specifications where savings is a categorical variable and 10% is the reference level of savings (see model 2 in table 14), and where savings is quadratic (see model 3 in table 14).

Results in model 2 suggest that moving from a 5% to a 10% savings level does not significantly increase the average individual utility. However, moving from 5% to 15% increases the utility by 3.639 points. On the other hand, we find no evidence of a quadratic effect of savings on utility. These results provide evidence to support the constant marginal returns hypothesis.

5.1.2 Second stage: random effects model estimates

Table 15 shows the results with the linear, categorical and polynomial savings specifications. In contrast to the first stage estimates, results in model 2 suggest that moving from a 5% to a 10% savings level significantly increases the likelihood to adopt a d-CEI by 15.6 pp, and moving from 5% to 15% increases the utility by 12.5 pp, which suggest a decreasing effect for every 5% increase. The quadratic specification yields a negative and highly significant coefficient (-0.94), thus supporting the hypothesis of decreasing marginal valuations over the savings attribute. This is an interesting result since it suggests a convex WTA threshold.

5.2 Analysis without individuals that fail the rationality test.

As mentioned in section 3.1, 7.6% of participants in the final sample failed the rationality test by choosing the dominated scenario (i.e. scenario B) in choice tasks 4 or 8 (see table 3). We conducted additional robustness checks to test how results may be affected by those individuals.

Table 16 shows the results of the part-worth utilities without the participant that failed the rationality test under different specifications for the savings attributes (i.e. linear, discrete and quadratic). Overall, the robustness checks are consistent with our main results in terms of the sign and significance. However, we find some differences in terms of the coefficient size. First, the coefficient corresponding to small health losses increases in magnitude in all three models. For instance, in the linear model the coefficient changes from -2.116 to -3.489. The same occurs for significant health losses: in the linear model we observe that the coefficient changed from -7.485 to -10.011. The reversibility and savings attributes are affected in a similar way. This is an expected result since irrational individuals will normally show a higher random component. Once they are removed from the analysis, the deterministic part increases and thus we obtain a bigger coefficient.

We also analyze the willingness to adopt in Table 17. The results suggest that the models excluding individuals who failed the rationality tests are consistent with our main results in terms of sign, size and significance of the coefficients.

6 Discussion

Our article sheds light on a set of interventions that have traditionally been neglected from both researchers and HTA agencies, as they are often perceived as “unethical” or “unacceptable”. This two-stage discrete choice experiment

contributes to the economic literature by identifying the preferences of health policy-makers towards d-CEIs and in particular to identify the importance of certain attributes in the decision to adopt such interventions. By doing so, our investigation provides a route for researchers and HTA professionals to head into the previously unexplored 'Far West'.

The study provides three main results. First, consistent with our hypothesis, we find that the respondents prioritize either the attribute on the health loss (as perceived by the patient population) or to the possibility to switch back to usual care (expressed in time) and cost-savings are considered secondary (least important attribute). Nonetheless, we observe a difference between the impact of attributes between the first stage (i.e. forced preference elicitation task) and the second stage (i.e. adoption preference task). It appears that the main attribute that affects utility is "significant health losses", while "reversibility (hardly possible)" appears to dominate over the rest. The result showing that reversibility plays a major role in adoption decisions compared to forced choices appears consistent with prospect theory and / or anticipated regret theory (Brewer et al. [2016]). Indeed, prospect theory posits greater loss aversion in the domain of losses. Therefore, decision-makers may be more risk averse in adoption preferences because d-CEIs imply health losses compared to the current situation (the status quo). In the conditional preference tasks, the status quo did not exist (forced choices). Another interpretation is that decision makers could anticipate regret in not being able to come back easily to the original situation (which is what the reversibility attribute captures). Unfortunately, we did not collect any information about the experience of potential regret to discriminate about one assumption or the other. Further analysis on the relationship between anticipated regret and adoption preferences in two-stages DCE tasks is warranted. Second, the stratified analysis of the compensation requirements according to their responses of the sensitivity attributes reveal the existence of preference heterogeneity across participants. Overall we find that compared to respondents who were willing to adopt, individuals who are not willing adopt the preferred d-CEI when disease severity increased to moderate or savings uncertainty increased to high have a higher compensation requirement to remain indifferent in a scenario when "health loss" changes from very low to significant. And third, the robustness checks suggest a linear effect of the savings attribute in the first stage, and a quadratic effect in the second. This finding suggests a constant marginal valuation of benefits in the first stage and, consistent with prospect theory, a diminishing sensitivity of benefits in the second stage.

There are several limitation in our study. First, the health attribute may be particularly difficult to understand for respondents. Health loss can be interpreted in terms of life duration, as quality of life or an interaction of both. As discussed by Attema et al. [2016], quality of life is hard to imagine, since people are more familiar with speaking about health in terms of ability to function in different activities, such as captured by the EQ-5D-5L classification system. Second, a non-negligible proportion of respondents is integrated by non-decision makers or HTA professionals. Third, although this discrete choice experiment allows to identify the importance of certain attributes in the adoption process, it was not designed to test the non-linearity hypothesis of the savings attribute: indeed, it is likely that beyond a certain threshold, savings no longer compensate

for the health loss or for the reduction in reversibility. Despite our findings that suggest a linear WTA in the first stage and a quadratic for the second, we only explore 3 different levels for the savings attributes and hence the DCE cannot provide conclusive evidence on the shape of the WTA. Finally, our design does not allow to capture interactions between attributes. This may be extremely relevant as respondents may not have the same preferences over attributes conditional on the value of other secondary attributes. Further research would be required to enquire these aspects.

An interesting aspect that was not discussed is into which extent the attributes estimates interact with other covariates such as country, age and gender, education, and occupation. Although not shown in the main results, we conducted the analysis and find no evidence of potential interactions, except education and occupation ⁵. Additionally, when analysing previous experience with d-CEIs among respondents we find more conclusive results. In particular, individuals agreeing to d-CEIs being considered as legitimate comparators in HTA committees had significantly lower utility losses associated with low health loss scenarios ⁶.

It is important to note that, as shown in the summary statistics section in table 2, we observe that some individuals (26%) claimed to have changed their opinion on d-CEIs during the course of the experiment. These results may suggest that making real trade-offs between health losses and savings requires familiarity with this type of reasoning. Individuals may be reluctant to consider d-CEIs at first, but providing information may allow individuals to better accept this type of intervention. In light of this, it might be interesting to further evaluate the effect of past experience with d-CEIs on willingness to adopt them. Finally, this study aims to identify the preferences of health policy-makers, but these preferences likely depend on the social acceptability of d-CEIs in the wider population. It would therefore be interesting to run this experiment in a wider population, beyond decision-makers.

From a policy point of view, d-CEIs are particularly relevant since the Covid-19 pandemic, which has potentially altered health policy-makers' preferences regarding these type of interventions. Indeed, some states have used d-CEIs, sometimes without naming them as such. This is the case with fabric masks, for example, and non-surgical masks which, for lack of available resources, have been readily adopted. Further research could explore into which extent COVID has affected policy makers preferences over d-CEIs.

⁵This could be explained by the fact that the sample is unbalanced for certain sub-categories of respondents.

⁶Results available upon request

References

- A. E. Attema, W. B. Brouwer, O. l'Haridon, and J. L. Pinto. An elicitation of utility for quality of life under prospect theory. *Journal of health economics*, 48:121–134, 2016.
- N. T. Brewer, J. T. DeFrank, and M. B. Gilkey. Anticipated regret and health behavior: A meta-analysis. *Health Psychology*, 35(11):1264, 2016.
- G. Calabrò, G. La Torre, C. De Waure, P. Villari, A. Federici, W. Ricciardi, and M. Specchia. Disinvestment in healthcare: an overview of hta agencies and organizations activities at european level. *BMC health services research*, 18(1):1–7, 2018.
- J. Cheng, E. Pullenayegum, D. A. Marshall, J. K. Marshall, and L. Thabane. An empirical comparison of methods for analyzing correlated data from a discrete choice survey to elicit patient preference for colorectal cancer screening. *BMC medical research methodology*, 12:1–17, 2012.
- D. A. Hensher, J. M. Rose, J. M. Rose, and W. H. Greene. *Applied choice analysis: a primer*. Cambridge university press, 2005.
- S. Hess and J. M. Rose. Can scale and coefficient heterogeneity be separated in random coefficients models? *Transportation*, 39(6):1225–1239, 2012.
- S. Hess and K. Train. Correlation and scale in mixed logit models. *Journal of choice modelling*, 23:1–8, 2017.
- D. Kahneman and A. Tversky. On the interpretation of intuitive probability: A reply to jonathan cohen. 1979.
- Z. Kallas and J. M. Gil. A dual response choice experiments (drce) design to assess rabbit meat preference in catalonia: A heteroscedastic extreme-value model. *British Food Journal*, 114(10):1394–1413, 2012.
- C. F. Manski. Daniel mcfadden and the econometric analysis of discrete choice. *The Scandinavian Journal of Economics*, 103(2):217–229, 2001.
- D. McFadden and K. Train. Mixed mnl models for discrete response. *Journal of applied Econometrics*, 15(5):447–470, 2000.
- D. McFadden et al. Conditional logit analysis of qualitative choice behavior. 1973.
- NICE. The nice methods of health technology evaluation: the case for change. 2020.
- D. Revelt and K. Train. Mixed logit with repeated choices: households' choices of appliance efficiency level. *Review of economics and statistics*, 80(4):647–657, 1998.
- A. Rotteveel, M. Lambooi, E. Over, J. Hernández, A. Suijkerbuijk, A. de Blaeij, G. de Wit, and N. Mouter. If you were a policymaker, which treatment would you disinvest? a participatory value evaluation on public preferences for active disinvestment of health care interventions in the netherlands. *Health Economics, Policy and Law*, pages 1–16, 2022a.

- A. H. Rotteveel, M. S. Lambooi, J. van Exel, and G. A. de Wit. To what extent do citizens support the disinvestment of healthcare interventions? an exploration of the support for four viewpoints on active disinvestment in the netherlands. *Social Science & Medicine*, 293:114662, 2022b.
- V. Soekhai, E. W. de Bekker-Grob, A. R. Ellis, and C. M. Vass. Discrete choice experiments in health economics: past, present and future. *Pharmacoeconomics*, 37:201–226, 2019.
- F. Traets, M. Meulders, and M. Vandebroek. Modelling consideration heterogeneity in a two-stage conjunctive model. *Journal of Mathematical Psychology*, 109:102687, 2022.

A DCE structure

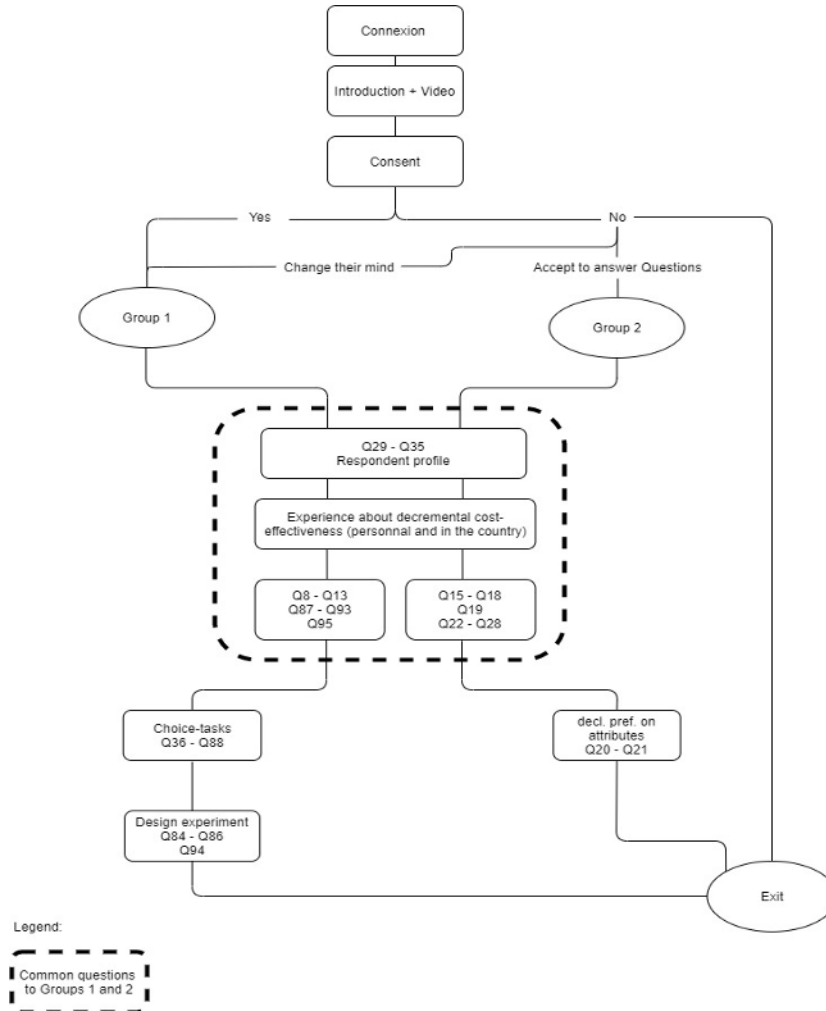


Figure 4: DCE structure

B About the experiment

The intervention relies on the following criteria being fulfilled and unchanged for all the choice sets:

- A reference treatment exists for the 10.000 patients suffering from a specific condition, and we refer to it as usual care. The two options presented in the choice-sets should be considered as decrementally cost-effective alternatives to usual care.
- The scenarios do not include patients experiencing end of life care.
- Disease severity of the patient (before treatment) is **LOW**
- Uncertainty related to the expected cost-savings is **LOW**

Figure 5: Context provided to participants before each choice task.

	OPTION 1	OPTION 2
HEALTH LOSS (as perceived by the patient population)	Very small	Significant
POSSIBILITY TO SWITCH BACK TO USUAL CARE (expressed in time)	Hardly possible due to long delays	Possible, at anytime
COST-SAVINGS (expressed in % of a fixed 100M€ healthcare budget)	5%	15%

Which option would you choose?

Option 1 Option 2

Would you be ready to substitute usual care by the option you selected?

Yes
 No

Figure 6: Example of choice task given to participant. This example corresponds to choice task number 2

Would you be ready to substitute usual care by the option you selected if disease severity of the patient (before treatment) changes to MODERATE?

- Yes
- No

Would you be ready to substitute usual care by the option you selected if uncertainty related to the expected cost-savings changes to HIGH?

Uncertainty indicates the possibility that the decremental cost-effective intervention won't bring the expected cost-saving.

- Yes
- No

Figure 7: Willingness to adopt sensibility attributes.

C Test of hypothesis comparing groups based on the sensitivity attributes

C.1 First-stage results

Attribute	2 vs 3	3 vs 4
Health loss: low	2.6	-2.3
Health loss: significant	7.0*	16.0*
Reversibility: possible with delay	3.9	5.5
Reversibility: hardly possible	5.6*	6.3

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 11: First stage test of hypothesis of coefficient differences across groups. The second column shows the p-values for the null hypothesis that attributes coefficients are the same in respondents who are and are not willing to substitute when disease severity increases from low to moderate; the third column displays the p-values for the null hypothesis that attributes coefficients are the same in respondents who are and are not willing to substitute when savings uncertainty increases to from low to high.

C.2 Second-stage results

Attribute	(2) vs (3)	(4) vs (5)
Health loss: low	-1.90	3.70
Health loss: significant	9.10*	15.90***
Reversibility: possible with delay	-9.50	-4.30
Reversibility: hardly possible	-8.50	0.70

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 12: Second stage test of hypothesis of coefficient differences across groups. The second column shows the p-values for the null hypothesis that attributes coefficients are the same in respondents who are and are not willing to substitute when disease severity increases from low to moderate; the third column displays the p-values for the null hypothesis that attributes coefficients are the same in respondents who are and are not willing to substitute when savings uncertainty increases to from low to high.

D Test of hypothesis of personal characteristics based on the sensitivity attributes

Variables	WTA if disease severity increases to moderate		WTA if savings uncertainty increases to high		Mean difference test	
	Yes: group 1	No: group 2	Yes: group 3	No: group 4	group 1 vs group 2	group 3 vs group 4
<35 yo	0.49	0.40	0.46	0.43	0.09	0.03
>35 yo	0.51	0.60	0.54	0.57	-0.09	-0.03
Female	0.56	0.55	0.53	0.56	0.01	-0.03
Male	0.44	0.45	0.47	0.44	-0.01	0.03
French	0.49	0.45	0.35	0.50	0.04	-0.15*
Non-French	0.51	0.55	0.65	0.50	-0.04	0.15*
HTA professional	0.18	0.19	0.23	0.17	-0.01	0.06
Government officer	0.07	0.06	0.08	0.06	0.01	0.02
Health economist	0.26	0.28	0.26	0.27	-0.02	-0.02
Student	0.26	0.23	0.21	0.25	0.03	-0.05
Clinician	0.10	0.06	0.08	0.07	0.04	0.00
Occupation: other	0.14	0.19	0.15	0.17	-0.05	-0.02
WTA if disease severity increases to high	1.00	0.00	0.63	0.34	1***	0.28***
WTA if high savings uncertainty	0.34	0.14	1.00	0.00	0.2***	1***

Table 13: Test of hypothesis of personal characteristics based on the sensitivity attributes. Column 2 and 3 show the mean values of respondents who are and are not willing to adopt the d-CEI if disease severity increases to moderate; Column 4 and 5 show the mean values of respondents who are and are not willing to adopt the d-CEI if savings uncertainty increases to moderate. Column 6 and 7 shows the differences between groups. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

E Testing non-linearities

E.1 First stage: mixed logit estimates

	Model 1 (linear savings)		Model 2 (categorical savings)		Model 3 (quadratic savings)	
	Mean (SE)	SD (SE)	Mean (SE)	SD (SE)	Mean (SE)	SD (SE)
Health loss: very small	<i>Ref</i>		<i>Ref</i>		<i>Ref</i>	
Health loss: small	-2.116*** (0.432)	1.255*** (0.377)	-1.697*** (0.402)	1.147*** (0.38)	-4.147*** (1.176)	2.463*** (0.78)
Health loss: significant	-7.845*** (1.039)	2.701*** (0.816)	-7.696*** (1.068)	3.445*** (0.628)	-11.305*** (2.22)	-3.513*** (0.714)
Reversibility: possible at any time	<i>Ref</i>		<i>Ref</i>		<i>Ref</i>	
Reversibility: with delay	-2.207*** (0.582)	1.084** (0.472)	-2.113*** (0.57)	0.82 (0.539)	-4.236*** (1.612)	1.241*** (0.415)
Reversibility: hardly possible	-4.027*** (0.561)	0.282 (0.639)	-3.972*** (0.555)	0.282 (0.639)	-5.848*** (1.249)	0.282 (0.639)
Savings: 5% of budget			<i>Ref</i>			
Savings: 10% of budget			1.211 (1.393)	2.109*** (0.772)		
Savings: 15% of budget			2.258*** (0.512)	-0.783* (0.469)		
Savings: linear 5% of budget	1.385*** (0.269)	-0.902*** (0.208)			-4.403 (9.037)	2.872 (3.426)
Savings: quadratic					1.629 (2.286)	0.183** (0.084)
N (individuals)	180		180		180	
Choice observations	1,408		1,408		1,408	
Log-likelihood	-549.53		-546.49		-547.21	

Standard errors in parenthesis
 *** p<0.01, ** p<0.05, * p<0.1.

Table 14: Comparison of econometric specifications for the “savings” attribute: mixed logit estimates

E.2 Second stage: random effects model estimates

	Linear savings	Categorical savings	Quadratic savings
	(1)	(2)	(3)
Health loss: very small	<i>Ref</i>	<i>Ref</i>	<i>Ref</i>
Health loss: small	-0.119*** (0.022)	-0.054* (0.030)	-0.054* (0.030)
Health loss: significant	-0.259*** (0.038)	-0.142*** (0.052)	-0.142*** (0.052)
Reversibility: always possible	<i>Ref</i>	<i>Ref</i>	<i>Ref</i>
Reversibility: with delay	-0.091*** (0.024)	-0.040 (0.028)	-0.040 (0.028)
Reversibility: hardly possible	-0.303*** (0.025)	-0.234*** (0.033)	-0.234*** (0.033)
Savings: linear 5% of budget	0.092*** (0.014)		0.437*** (0.105)
Savings: 5% of budget		<i>Ref</i>	
Savings: 10% of budget		0.156*** (0.024)	
Savings: 15% of budget		0.125*** (0.033)	
Savings: quadratic			-0.094*** (0.028)
N(Individuals)	180	180	180
Choice observations	1408	1408	1408

Standard errors in parenthesis *** p<0.01, ** p<0.05, * p<0.1.

Table 15: Comparison of econometric specifications for the “savings” attribute: random effects model estimates

F Results without irrational individuals.

F.1 First stage: mixed logit estimates

	Model 1 (linear savings)		Model 2 (categorical savings)		Model 3 (quadratic savings)	
	Mean (SE)	SD (SE)	Mean (SE)	SD (SE)	Mean (SE)	SD (SE)
Health loss: very small	<i>Ref</i>		<i>Ref</i>		<i>Ref</i>	
Health loss: small	-3.489*** (0.912)	0.088 (0.623)	-3.664*** (1.064)	-0.334 (0.711)	-3.922*** (1.172)	0.339 (0.722)
Health loss: significant	-10.011*** (1.847)	-0.116 (0.710)	-10.086*** (2.218)	3.473*** (0.916)	-10.777*** <i>Ref</i>	3.182*** (0.933)
Reversibility: always possible	<i>Ref</i>		<i>Ref</i>		<i>Ref</i>	
Reversibility: with delay	-4.347*** (1.288)	1.770*** (0.528)	-3.866*** (1.382)	1.371* (0.771)	-4.260** (1.684)	1.537** (0.650)
Reversibility: hardly possible	-5.208*** (1.033)	1.137** (0.478)	-5.148*** (1.128)	-0.743 (0.654)	-5.476*** (1.276)	0.585 (0.721)
Savings: 5% of budget			<i>Ref</i>			
Savings: 10% of budget			3.818 (3.362)	0.573 (0.926)		
Savings: 15% of budget			3.639*** (1.017)	0.060 (0.771)		
Savings: linear 5% of budget	1.588*** (0.374)	-0.767** (0.356)			9.205 (12.344)	0.573 (0.926)
Savings: quadratic					-1.824 (3.063)	0.060
N (individuals)	166		166		166	
Choice observations	1300		1300		1300	
Log-likelihood	-427.16		-426.33		-424.18	

Standard errors in parenthesis
 *** p<0.01, ** p<0.05, * p<0.1.

Table 16: First stage results without individuals who failed the rationality test in choice task 4 or 8.

F.2 Second stage: random and fixed effects model estimates

	Linear Savings		Categorical Savings		Quadratic savings	
	(1)	(2)	(3)	(4)	(5)	(6)
Health loss: very small	<i>Ref</i>	<i>Ref</i>	<i>Ref</i>	<i>Ref</i>	<i>Ref</i>	<i>Ref</i>
Health loss: small	-0.117*** (0.023)	-0.114*** (0.023)	-0.050 (0.032)	-0.048 (0.032)	-0.050 (0.032)	-0.048 (0.032)
Health loss: significant	-0.247*** (0.041)	-0.240*** (0.042)	-0.127** (0.058)	-0.121** (0.059)	-0.127** (0.058)	-0.121** (0.059)
Reversibility: always possible	<i>Ref</i>	<i>Ref</i>	<i>Ref</i>	<i>Ref</i>	<i>Ref</i>	<i>Ref</i>
Reversibility: with delay	-0.090*** (0.024)	-0.087*** (0.024)	-0.036 (0.031)	-0.034 (0.030)	-0.036 (0.031)	-0.034 (0.030)
Reversibility: hardly possible	-0.318*** (0.026)	-0.322*** (0.026)	-0.248*** (0.035)	-0.253*** (0.035)	-0.248*** (0.035)	-0.253*** (0.035)
Savings: linear 5% of budget	0.094*** (0.015)	0.088*** (0.015)			0.427*** (0.116)	0.416*** (0.115)
Savings: 5% of budget			<i>Ref</i>	<i>Ref</i>		
Savings: 10% of budget			0.154*** (0.025)	0.147*** (0.025)		
Savings: 15% of budget			0.126*** (0.036)	0.114*** (0.036)		
Savings: quadratic					-0.091*** (0.031)	-0.090*** (0.031)
Constant	0.779*** (0.031)	0.789*** (0.026)	0.792*** (0.038)	0.797*** (0.034)	0.456*** (0.116)	0.471*** (0.114)
Choice observations	1300	1300	1300	1300	1300	1300
N(Individuals)	166	166	166	166	166	166
Individual Fixed Effect	No	Yes	No	Yes	No	Yes

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 17: Second stage results without individuals who failed the rationality test in choice task 4 or 8.