Fairness in cost-benefit analysis: a methodology for health technology assessment

26th July 2016

Anne-Laure Samson, Erik Schokkaert, Clémence Thébaut, Brigitte Dormont, Marc Fleurbaey, Stéphane Luchini, Carine Van de Voorde[1][2]

ABSTRACT. We evaluate the introduction of various forms of antihypertensive treatments in France with a distribution-sensitive cost-benefit analysis. Compared to traditional cost-benefit analysis, we implement distributional weighting based on equivalent incomes, a new concept of individual well-being that does respect individual preferences but is not subjectively welfarist. Individual preferences are estimated on the basis of a contingent valuation question, introduced into a representative survey of the French population. Compared to traditional cost-effectiveness analysis in health technology assessment, we show that it is feasible to go beyond a narrow evaluation of health outcomes while still fully exploiting the sophistication of medical information. Sensitivity analysis illustrates the relevancy of this richer welfare framework, the importance of the distinction between an ex ante and an ex post-approach, and the need to consider distributional effects in a broader institutional setting.

[1]ALS: PSL, Université Paris-Dauphine; ES: Department of Economics, University of Leuven and CORE, Université catholique de Louvain; CT: Université de Limoges, OMIJ, IAE Limoges; BD: PSL, Université Paris-Dauphine; MF: Princeton University; SL: Aix-Marseille University (Aix-Marseille School of Economics), CNRS and EHESS; CVDV: Department of Economics, University of Leuven.

[2]The research and the survey used in this paper were partly funded by the Health Chair which is financed by MGEN, a non-profit complementary health insurance organization for teachers in France, and by ISTYA, a federation of non profit complementary health insurance organizations for French civil servants. The rules monitoring the sponsorship by MGEN and ISTYA prevent any interference with research projects. In particular, the sponsors did not have any influence in study design; in the collection, analysis and interpretation of data; in the writing of the report, as well as in the decision to submit this paper.

1 Introduction

In many countries with a publicly financed health care system, governments are concerned about the increase in health care expenditures. For a welfare economist the most natural approach to evaluate these expenditures is cost-benefit analysis (CBA). Yet, both the medical community and most health economists show a deep reluctance against traditional CBA, i.e. the approach in which maximization of the unweighted sum of individual consumer surpluses (or related concepts such as compensating variations) is taken as the social objective. Using willingness-to-pay as a monetary valuation of health changes is considered morally objectionable. Moreover, the fact that the rich can have a higher willingness-to-pay for a treatment just because they can afford it is seen as inequitable. As an alternative, it has become common practice to implement cost-effectiveness analysis (CEA), using as objective function the sum of health outcomes (e.g. as measured in QALY's). This is sometimes advocated as a non- or extra-welfarist approach.

While CEA is useful in guiding choices within a given health care budget, it has obvious limitations. First, focusing exclusively on health outcomes is not sufficient to determine the optimal size of the health care budget since this requires an analysis of the trade-off between health and other dimensions of life. Second, although in theory it is possible to introduce distributional (e.g. severity) weighting in CEA, the usual practice is to take as social objective the unweighted sum of health outcomes. This approach is not satisfactory from an ethical perspective. Many find it difficult to accept that richer citizens can pay out of pocket for therapies that are insufficiently cost-effective and that are therefore denied to the poor. More fundamentally, equity requires comparing individuals in terms of all the relevant dimensions of life, and not just in terms of health. "A state of affairs in which those who are otherwise worse off are healthier than those who are otherwise more fortunate is more just rather than less just than a state of affairs which is exactly the same except that health is equally distributed" (Hausman, 2007).

In this paper we show that it is possible to formulate a version of CBA which answers the criticism raised by health economists and does not suffer from the limitations of CEA. More specifically we propose to perform economic evaluation in health care with a *distributionally sensitive* social welfare function of individual *well-being* levels in an *ex post*-perspective. First, we implement as a measure of individual well-being the so-called

equivalent income. This measure does respect individual preferences, but does not coincide with subjective utility. Second, we will introduce distributional weights through a social welfare function (SWF). This is not innovative from a welfare economic perspective, since it has been long accepted that the traditional approach without distributional weighting is logically flawed and ethically unattractive (see, e.g., Blackorby and Donaldson, 1990; Fleurbaey and Blanchet, 2013). Third, both CEA and traditional CBA handle uncertainty from an *ex ante* perspective, i.e. by first calculating expected values of the outcomes (respectively health or utility) at the individual level and then defining the social welfare function over these expected values. We will advocate an *ex post* perspective, which takes into account the inequality in the final outcomes in each of the possible social states. In this paper we do not discuss deeply the ethical justification for each of these choices (see, e.g., Fleurbaey et al., 2013). Our main contribution is to show that this rich approach can be applied in a real-world setting. Our empirical application is the assessment of antihypertensive treatments in France.

Section 2 presents our theoretical background in general terms. Section 3 describes the decision problem, the data and the empirical procedure. Results are presented in section 4. Section 5 concludes.

2 Theoretical background

As soon as one aims to go beyond health, it becomes essential to formulate assumptions about the desirable trade-offs between health and other dimensions of life. To keep the empirical analysis tractable, we focus in this paper on only two dimensions, income y and health h, and we describe the life situation of individual i by the bundle (y_i, h_i). In a welfarist approach the ultimate criterion to evaluate the individual's situation is their level of subjective satisfaction. In section 2.1 we propose an alternative welfare measure, the so-called equivalent income. We then argue that distributional considerations can be modelled in a natural way through a social welfare function (section 2.2). We finally discuss the distinction between an *ex ante* and an *ex post* approach in a context with uncertainty (section 2.3).

2.1 A welfare measure: equivalent income

We assume that individual i has preferences defined over bundles (y_i, h_i), capturing what he considers to be important in life, i.e. his own personal life project.[3] We write $(y_i, h_i)R_i(y_i^{'}, h_i^{'})$ if individual i weakly prefers (y_i, h_i) to $(y_i^{'}, h_i^{'})$ (with I_i indicating indifference). As we believe that each individual is best placed to decide about the trade-offs between the various dimensions of his/her own life, we look for an individual welfare measure $v_i(y_i, h_i)$ that does respect these individual ideas about a good life, i.e. such that

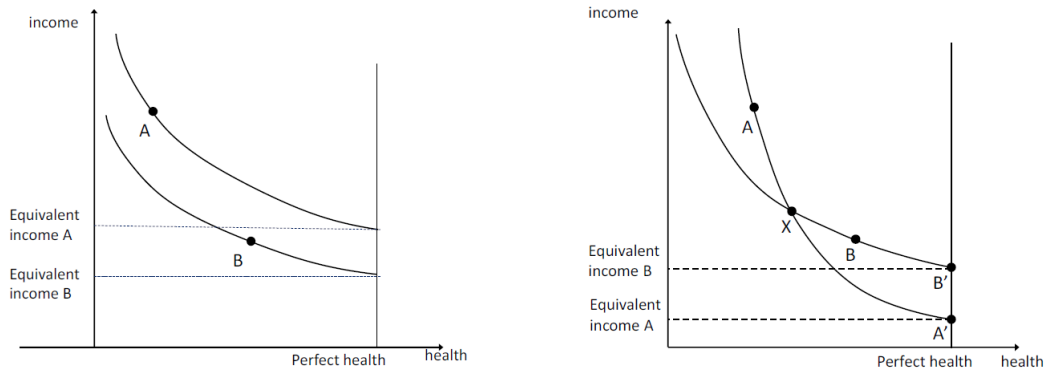
$$v_i(y_i^{'}, h_i^{'}) \geq v_i(y_i, h_i) \Leftrightarrow (y_i^{'}, h_i^{'})R_i(y_i, h_i). \tag{1}$$

Eq. (1) shows that $v_i(y_i, h_i)$ can be interpreted as an individual utility function. Yet imposing (1) does not boil down to subjective welfarism. There are many different functions that respect ordinal preferences, all positive monotonic transformations of each other. Taking subjective well-being or happiness is just one possible choice out of all these possible cardinalisations. The problem with this specific happiness-scale becomes obvious when one extends the idea of respecting individual preferences to interpersonal comparisons. Take two individuals i and j with identical preferences, say $R_i = R_j = \overline{R}$. Respecting these common preferences then imposes

$$v_i(y_i^{'}, h_i^{'}) \geq v_j(y_j, h_j) \Leftrightarrow (y_i^{'}, h_i^{'})\overline{R}(y_j, h_j). \tag{2}$$

Figure 1a illustrates. Ann (in A) and Bob (in B) have identical preferences. Eq. (2) then implies that Ann is better off than B, since A is on a higher indifference curve. Yet this does *not* imply that the subjective satisfaction of Ann is larger than that of Bob. Introducing the notation $S_{Ann}(X)$ for the subjective satisfaction of A in X, respecting ordinal preferences implies in our example that $S_{Ann}(A) > S_{Ann}(B)$ and that $S_{Bob}(A) > S_{Bob}(B)$. Yet this does not preclude that $S_{Bob}(B) > S_{Ann}(A)$, which would go against condition (2).

[3] "Preferences" have been interpreted in many different ways in the economic literature. The most popular interpretation refers to "revealed preferences", i.e. the preferences that can be derived from individual rational choices. This is *not* our interpretation. We believe that choice behaviour does not always reveal the true underlying convictions of individuals, e.g. because of informational and decision-making limitations. The preferences, as we define them, are a mental construct, a representation of the (cognitive) life project of the individual.



a. Identical preferences b. Different preferences

Figure 1: Equivalent income

It is possible that both individuals prefer A to B, while at the same time the individual in B is more satisfied than the individual in A. This can occur, for example, if individuals adapt their aspirations to what is feasible for them in their actual situation.[4]

The challenge is then to formulate a concept of individual well-being that does respect condition (2) and can be applied if individuals have different preferences. Here is one possibility.[5] Take "perfect health" h^* as a reference level for health and define the equivalent income y_i^* of individual i implicitly by

$$(y_i, h_i) I_i (y_i^*, h^*), \tag{3}$$

i.e. the hypothetical income that together with perfect health would put the individual in a situation that is for her as good as her actual situation (y_i, h_i). The proposal is to take this equivalent income as the definition of individual well-being, i.e.

$$v_i(y_i, h_i) \equiv y_i^*(y_i, h_i). \tag{4}$$

The intuition behind this proposal can be explained with Figure 1b. For Ann, her actual

[4]This was called "physical-condition neglect" in Sen (1985). The empirical happiness literature has collected a lot of information showing that this phenomenon also occurs in the case of health (see, e.g., Loewenstein and Ubel, 2008).

[5]A more elaborate justification of the concept of equivalent income and a comparison with alternative ways of measuring well-being can be found in Fleurbaey and Blanchet (2013) and in Decancq et al. (2015).

situation A is equivalent to the hypothetical situation A' (in which she is in perfect health). Similarly, for Bob his actual situation B is equally good as the hypothetical situation B'. The assumption underlying (4) is that we can compare the well-being levels of Ann and Bob in A' and B' on the basis of the incomes in these situations, i.e. on the basis of the equivalent incomes indicated in the figure. To see why this is meaningful, consider another hypothetical situation X, in which Ann and Bob have the same health and the same income. If we take preferences into account, the well-being of Ann will be lower than that of Bob, since the shape of the indifference curves indicates that she cares more about being ill. However, if Ann were in situation A' and Bob were in situation B', there is no welfare loss due to ill-health. It seems natural to assume that in such a situation preference differences should not count, and since health is the same in A' and B', the comparison can then be fully based on the incomes in these situations. Moreover, if we can use "equivalent incomes" as a measure of well-being to compare A' and B', we can use them also to compare A (equivalent to A') and B (equivalent to B').

The equivalent income approach is *not* subjectively welfarist. Figure 1a immediately shows that, if two individuals share the same preferences, the one that reaches a higher indifference curve will always have a larger equivalent income. Moreover, it does not fall into the opposite trap of money fetishism either. Despite the fact that it is expressed in monetary terms (with the ensuing practical advantages), it is an encompassing measure of well-being, taking into account the welfare loss as a consequence of being ill. Eq. (3) can be written as

$$y_i^*(y_i, h_i) = y_i - WTP_i(h_i \rightarrow h^*), \tag{5}$$

where $WTP_i(h_i \rightarrow h^*)$ denotes the willingness-to-pay of individual i to go from his actual health situation h_i to the "perfect" health level h^*. This willingness-to-pay can be large, and the ranking of individuals on the basis of equivalent incomes can be very different from their ranking in terms of income. In Figure 1b, Ann's income in A is larger than Bob's income in B, but nevertheless her equivalent income is lower (because her health situation is worse, and she cares greatly about health).

2.2 Distribution: the social welfare function

Suppose there are N individuals. If the ultimate criterion to evaluate social states and policies is the well-being of the individuals, all the relevant information is contained in the vector $(y_1^*, \ldots, y_N^*)$. For the moment we keep N fixed, and we will explain later how we handle mortality in our empirical work. Uncertainty will be introduced in the next section.

Denoting the equivalent income of individual i after policy A by y_i^{*A}, policy A is better than policy B if $(y_1^{*A}, \ldots, y_N^{*A})$ is socially preferred to $(y_1^{*B}, \ldots, y_N^{*B})$. We can represent this social preference relation by a social welfare function $W(y_1^*, \ldots, y_N^*)$, with the functional form of $W(.)$ capturing the specific stance on (re)distribution. For an inequality-averse social welfare function it holds that $\partial W/\partial y_i^* > \partial W/\partial y_j^*$ if $y_i^* < y_j^*$. This takes care of the equity problem related to the use of willingness-to-pay. As an illustration, take two individuals with $y_i < y_j$ but with the same monetary value of the welfare loss due to illness, i.e. $WTP_i(h_i \to h^*) = WTP_j(h_j \to h^*)$. Since individual j can better afford to pay for a better health, it may be misleading to infer from the equality of the willingnesses-to-pay that they both care equally about being ill. This is taken into account in the inequality-averse SWF, however. Eq. (5) shows that in this example $y_i^* < y_j^*$ and therefore the same individual loss will have a smaller negative effect on social welfare for the rich individual j than for the poor individual i.[6]

In our empirical work, we will use the popular iso-elastic Atkinson SWF, defined as

$$SW = \frac{1}{1-\rho} \sum_i (y_i^*)^{1-\rho}, \tag{6}$$

with ρ the parameter of inequality aversion. If one is not averse towards inequality, $\rho = 0$, and eq. (6) becomes the simple sum of equivalent incomes. If ρ increases, a relatively larger and larger weight is given to the worse-off individuals. In the extreme case where $\rho \to \infty$, eq. (6) boils down to the maximin social welfare function, which gives a positive

[6] It is well-known that choosing either money or health as the numeraire does matter in an approach without distributional weights. The choice becomes irrelevant, however, as soon as one works within a coherent SWF-framework - see, e.g., the exchange of ideas between Brekke (1997), Drèze (1998) and Johansson (1998).

weight to only the worst-off individuals.[7] The choice of ρ is a value judgment. In our empirical work we will show the results for different values of ρ.

2.3 Uncertainty: ex ante versus ex post

Most health care interventions have uncertain outcomes. We will therefore now assume that there are S different possible states of the world, occurring with probabilities $(\Pi_1,\ldots,\Pi_S)$ with $\sum_s \Pi_s = 1$. We denote the vector of equivalent incomes in state s by $(y^*_{1s},\ldots,y^*_{Ns})$.

The *ex ante*-approach first computes expected outcomes at the individual level

$$EY^*_i = \sum_s \Pi_s u_i\left(y^*_{is}\right), \tag{7}$$

with $u_i(y^*_{is})$ a von Neumann-Morgenstern utility function, and then introduces these individual expected outcomes into the social welfare function. Under the assumption of constant relative risk aversion ε, we can write $u_i(y^*_{is})$ as

$$u_i\left(y^*_{is}\right) = \frac{1}{1-\varepsilon}\left(y^*_{is}\right)^{1-\varepsilon}. \tag{8}$$

In the Atkinson-specification, one gets

$$SW^{exante} = \frac{1}{1-\rho}\sum_i \left(EY^*_i\right)^{1-\rho} = \frac{1}{1-\rho}\sum_i \left[\sum_s \Pi_s u_i\left(y^*_{is}\right)\right]^{1-\rho}. \tag{9}$$

The inequality aversion pertains to individual expected outcomes.

In the *ex post*-approach one introduces inequality aversion with respect to the actual outcomes in each potential state of the world, and then computes the expected value of the

[7] For marginal changes, eq. (6) implies that the "distributional weight" of individual i is given by $(y^*_i)^{-\rho}$. An overview of such marginal weights with equivalent incomes calculated from a survey in Marseille is given in Fleurbaey et al. (2013). However, in the empirical application of this paper we will consider discrete changes and we will therefore directly compare different values of SW.

social welfare function. This yields

$$SW^{ex\,post} = \sum_s \Pi_s \left[\frac{1}{1-\rho} \sum_i (y_{is}^*)^{1-\rho} \right] \tag{10}$$

In the case considered in this paper, however, there is an important simplification. The risks we consider are health risks at the individual level. Since each individual's risk is independent of other individuals' risks, there is no macrorisk. In that case, by the law of large numbers, the final distribution of individual situations is almost certain; i.e. the term in brackets is almost always the same independently of s.[8] Equation (6) then remains the relevant definition of social welfare at the social level.

individual state	S1	S2
individual outcome	5	10

Policy A

individual state	S1	S2
individual outcome	7.5	7.5

Policy B

Table 1: Ex ante versus ex post

It is clear that the functions (9) and (10) coincide if individuals have no risk aversion and if society does not care about inequality, i.e., if $\varepsilon = \rho = 0$.[9] In general, however, the functions will be different. Compare in Table 1 the outcomes of policies A and B with a large number of individuals and two possible states for each individual, each occurring independently for every individual with a probability of 0.5.[10] In an ex ante-perspective, policies A and B are equivalent, as they yield the same vector of (equal) expected outcomes. From an ex post-perspective it matters that outcomes are equally distributed in all states with policy B, while with policy A, almost surely, half of the individuals will have outcome 5 and half will have outcome 10. If we want to give a greater weight to those that are "unlucky" in the different social states, one should favour an ex post-approach.[11]

[8] In the case of macrorisk (e.g., a pandemic), formula (10) can also have a social von Neumann-Morgenstern function representing risk attitude at the social level: $\sum_s \pi_s V\left[EDE(y_{1s}^*, ..., y_{Ns}^*)\right]$, where V embodies the social attitude to risk and EDE_s is the equally distributed equivalent in state s, i.e. $\sum_i (y_{is}^*)^{1-\rho} = N.(EDE_s)^{1-\rho}$ (see, e.g., Fleurbaey and Zuber, 2015).

[9] A special case is that of traditional cost-effectiveness analysis with the unweighted sum of QALYs as the objective function. In that setting there is no difference between ex ante and ex post evaluation.

[10] We therefore have 2^n social states s.

[11] See, e.g., Fleurbaey (2008) for a more general discussion of the pros and cons of ex ante versus ex post social evaluation in a context of risk.

However, we will also show the results of the ex ante-approach in our sensitivity analysis in section 4.2.1.

3 Empirical procedure

3.1 The problem: evaluating antihypertensive treatments

Our empirical application is the assessment of three different treatments for patients with essential hypertension in France, i.e. patients with high blood pressure (over 150 mmHg) but without a history of cardiovascular events. Prescribing antihypertensive treatment to these patients aims at controlling arterial blood pressure and therefore at decreasing the probability of occurrence of cardiovascular events (angina, myocardial infarction, stroke, heart failure), renal failure and end-stage renal failure. We focus on the comparison of three strategies:

- Strategy A is the placebo comparator.[12] Patients are not treated with any antihypertensive treatment in primary prevention. Physicians only start prescribing an active antihypertensive treatment after patients have experienced an event.
- With Strategy B, every patient is treated with ACE inhibitors in first-line treatment, with a bitherapy combining ACE inhibitors-diuretics in second-line treatment and finally with a tritherapy in third-line treatment. This is the cheapest strategy when both the cost of the treatment and the cost of avoided medical care are taken into account (HAS, 2012).
- With Strategy C, every patient is treated with calcium antagonists in first-line treatment, with a bitherapy combining calcium antagonists-ACE inhibitors in second-line treatment and with tritherapy in third-line treatment. This is the most effective strategy in terms of life years gained (HAS, 2012).

[12]The term "placebo comparator" is a misnomer, since patients do not receive a "fake drug", there is hence no placebo effect. We could also have called it a "do nothing"-strategy.

We use the model produced for HAS by IMS Health to calculate the costs of these three strategies and the resulting risks of cardiovascular events and renal failure disease. Throughout our calculations, we work with a time horizon of 10 years.

3.2 Data and the estimation of preferences

When working with distributional weights, it is not sufficient to collect information on a sample of patients that suffer from hypertension. Not only the total cost of the treatment, but also the distribution of that cost over the population has to be taken into account. We have therefore collected the necessary data from a representative sample of the French population. 3,331 individuals were interviewed in 2009 using computer-assisted face-to-face interviews.

The survey contains the usual questions on demographic and socioeconomic characteristics, detailed questions on specific diseases and health problems that the respondent might have experienced in the previous twelve months and a question on overall self-assessed health (SAH), where use was made of a visual 0-100 scale. The respondent was also asked about her lifestyles (smoking habits, alcohol consumption, weight and height, ...). After respondents had considered their own economic and health situation, they were confronted with a retrospective willingness-to-pay question (meant to measure $WTP_i(h_i \rightarrow h^*)$) about the exact amount of income they would have been willing to give up in exchange for having been in a state of perfect health during the last twelve months. Because some respondents refused to answer this question, our final sample consists of 2413 individuals.[13]

Basic features of our data are presented in table 2, where individuals with no hypertension are distinguished from individuals with hypertension and no cardiovascular event. This latter (crucial) group represents 15.7% of our original sample. This figure is lower than the prevalence in the French population, which is likely to be larger than 30%.[14] This un-

[13] More information (including the original formulation of the questions in French) can be found in Fleurbaey et al. (2012) and in the online appendix.

[14] 31% according to the "Étude Nationale Nutrition Santé" (Godet-Thobie H., et al., 2008); 47% of men and 35% of women in the French population aged between 35-74 years old according to data from the MONA LISA cohort (Wagner et al., 2011).

	All individuals	Individuals with no hypertension	Individuals with hypertension
Share of women	51.3	46.7	61.3
Age	48.5 (18.4)	44.1 (17.6)	58.1 (16.3)
SAH	72.3 (18.5)	74.9 (18.6)	66.6 (17.1)
Personal Income	1,247.2 €(690.7 €)	1,256.2 €(720.3 €)	1,227.2 €(620.9 €)
Equivalised Income	1,341.7 €(894.6 €)	1,348.6 €(968.9 €)	1,326.5 €(704.5 €)
WTP	69.9 €(169 €)	62.2 €(158 €)	86.9 €(190.2 €)
Equivalent Income	1,271.8 €(881.6 €)	1,293.3 €(973.8 €)	1,239.6 €(687.6 €)
Nb of obs.	2,413	2,035 (84.3%)	378 (15.7%)

Table 2: Descriptive information

derestimation of hypertension is a common phenomenon in surveys, as many individuals are not aware that they suffer from hypertension. All results are therefore weighted with sample weights based on age, gender and the prevalence of hypertension. The average willingness-to-pay for being in perfect health is 62.2€ for individuals with no hypertension. Reassuringly, in the subsample of individuals with hypertension, this willingness-to-pay is larger (86.9€). The low average willingness-to-pay can be explained by the relatively low incomes and high values of SAH in our sample. It can also be due to the fact that the WTP question referred to the health situation in the *previous* twelve months. Mortality issues were excluded and so was the anxiety related to uncertainty about future health.

Using eq. (5), we computed equivalent incomes for all individuals in the sample. The mean and median equivalent incomes are €1271.8 and €1150.0 respectively. To simulate the effect of different policies on equivalent incomes, we needed, in addition, information on individual preferences. The methodology used to estimate these preferences can be found in Schokkaert et al. (2013).

3.3 Evaluation of the three strategies[15]

3.3.1 The distribution of events resulting from the different strategies

We used the HAS prediction model to define the possible health trajectories that an individual may experience during a period of ten years. Individuals with hypertension can experience different cardiovascular events (stroke, angina, myocardial infarctus, heart failure) as well as renal failure and end-stage renal failure. They can die as a consequence of this event or they can survive. To keep the problem tractable we introduced some simplifying assumptions. Only one event can occur in each year; individuals may experience only two events during the 10-year period; and each event happens in the beginning of the year. Even with these assumptions, each individual can follow 3376 different trajectories.

We then calculated with the HAS model for each individual the probability of following any of the possible trajectories on the 10-year horizon. This probability depends on the chosen treatment strategy, but also on the kind of event, on the timing of this event (first or second event) and on the individual's characteristics (gender, age, diabetes and smoking habits). Since all individuals in the sample, with or without hypertension, may die from other causes, we implemented an "all causes mortality rate" dependent upon age and gender.[16]

Denote the probability that individual i follows the specific trajectory (path) p by π_{ip}, with $\sum_p \pi_{ip} = 1$ for all i. In the ex post-approach, these ex ante-probabilities at the individual level are interpreted as ex post-shares of the population, i.e. we will assume that a fraction π_{ip}/N of the population follows path p. Each of the possible situations on a path will result in specific health and income levels that do not only depend on the event experienced in the last period but also on the history, i.e., on the specific trajectory that brought the individual to that state, and on his starting position in period 0.

[15] All estimation results are presented extensively in the online appendix.

[16] We used the mortality rates that were produced for 2009 by the French Institute for Demographic Studies (INED).

3.3.2 Simulating the effects of events on health

First, we estimate the effect of the different events on health as measured by the SAH variable with a simple linear regression:

$$SAH_{i,0} = c + \sum_j s_j ev_{ij,0} + a_1 age_{i,0} + a_2 age_{i,0}^2 + \sum_k b_k x_{ik,0} + u_{i,0}, \tag{11}$$

where the subscript 0 indicates that the regression is run on our sample data (i.e. each individual is in period 0 of his 10-year trajectory), $ev_{ij,0} = 1$ if individual i experienced event j in the previous 12 months, $age_{i,0}$ is the age of individual i at the time of the interview, $x_{ik,0}$ indicates a list of control variables and $u_{i,0}$ is a disturbance term.

We then simulate the dynamic development of SAH over time (for $\tau = 1, \ldots, 10$) for an individual that is still alive in period τ:

$$SAH_{i,\tau} = SAH_{i,\tau-1} + \sum_j \hat{s}_j ev_{ij,\tau} + \left[\hat{a}_1 + \hat{a}_2 (age_{i,\tau}^2 - age_{i,\tau-1}^2)\right], \tag{12}$$

where the hats indicate estimated coefficients. Eq. (12) implements a series of assumptions. First, all individuals undergo a natural depreciation of health as a function of age, whether they suffer from hypertension or not. This is captured by the terms in square brackets. Second, when individual i experiences a cardiovascular event at the beginning of period τ (i.e. if $ev_{ij,\tau} = 1$), this has a negative effect on her self-assessed health as measured by $\hat{s}_j$.[17] Third, the socio-demographic control variables (including lifestyle) are kept constant throughout the 10-year period. Fourth, when the individual dies in period τ, be it as the consequence of a cardiovascular event or from other causes, we assume $SAH_{it} = 0$ for the periods $t = \tau, \ldots, 10$.

The resulting development of *SAH* will depend on the trajectory followed by the individual. We denote by $SAH_{ip\tau}$ the self-assessed health reached in period τ by individual i when she follows trajectory p.

[17] This means that we assume that the impact of an event on SAH is permanent, and that the impacts of different events are additive.

3.3.3 Simulating the effects of events on income

As far as income is concerned, we must distinguish two types of effects. First, income changes over time may be affected by the occurrence of cardiovascular events. Second, the costs of medical care must be allocated to the different individuals in the sample.

The change of income over time We assume that health events only influence the incomes of working individuals. The income development for the other groups of the population is simulated following the French social security regulation (see online appendix). For the subsample of working individuals who are less than 65 years old (the age of retirement), we estimate the following regression

$$y_{i,0} = d + hSAH_{i,0} + l.age_{i,0} + \sum_k m_k z_{ik,0} + v_{i,0}, \tag{13}$$

with $y_{i,0}$ indicating the level of monthly personal income, $z_{ik,0}$ the control variables and $v_{i,0}$ a disturbance term.[18]

Their income in period τ is calculated as

$$y_{i,\tau} = y_{i,\tau-1} + \hat{h}(SAH_{i,\tau} - SAH_{i,\tau-1}) + \hat{l}. \tag{14}$$

The coefficient $\hat{l}$ captures the appreciation of income with age. The value of $SAH_{i,\tau}$ is first simulated using eq. (12) and then introduced into eq. (14).

Finally, to calculate equivalized incomes, we assume that the family composition and the income of the partner (if there is one) do not change over the 10-year period. We denote by $y_{ip\tau}$ the equivalized income obtained in period τ by individual i when he follows trajectory p.

[18] Other functional forms (taking a logarithmic transformation of income as the dependent variable and/or introducing age squared as an explanatory variable) did not give a better fit than the simple linear form (13). Of course, this estimation on cross sectional data does not at all allow for a causal interpretation. There is a huge body of literature showing convincingly that a simple estimation of (13) will suffer from simultaneity bias. Our results should only be seen as a first illustration of how the association between health and income can be taken into account in a richer approach to policy evaluation.

Allocating the costs of medical treatment The costs of antihypertensive treatment (based on the costs of the drugs prescribed in the different strategies), the costs of the follow-up of the treatment (physician consultations, measurement of blood pressure, laboratory testing, etc.) and the medical costs induced by the occurrence of a cardiovascular event (costs of hospitalisation and of follow-up treatment) are all taken from the HAS model. They are classified into two categories.

First, the out-of-pocket payments (OOP) are calculated according to the French health insurance regulation and subtracted from the income of the concerned individuals. Second, the costs that are borne by the national health insurance are calculated for each period τ as

$$TC_\tau = \sum_i \sum_p \pi_{ip} C_{ip\tau}, \tag{15}$$

where π_{ip} is the probability that individual i follows trajectory p and $C_{ip\tau}$ is the corresponding cost (after subtracting OOP-payments) in period τ. Remember that in the ex post-approach the probabilities π_{ip} must be interpreted as population shares. This total cost TC_τ is divided over all the individuals in the sample, i.e. over all tax payers, proportionally to their actual contribution to the system, which in turn is determined by their (simulated) equivalized income $y_{ip\tau}$ on the basis of the tables in Caussat et al. (2005).[19]

3.3.4 Estimating the total welfare change for society

We can now calculate for all individuals the income and health levels that correspond to each possible state in their potential trajectories, differentiated for the three strategies considered. We then simulate their equivalent incomes $y^*_{ip\tau}$ and insert these values into eq. (6) to calculate social welfare. In our baseline simulation, we do this period by period. Adapting eq. (6) to our framework gives

$$SW_\tau^{ex\,post} = \frac{1}{1-\rho} \sum_p \sum_i \pi_{ip} (y^*_{ip\tau})^{1-\rho} \tag{16}$$

[19] In dividing the total costs over all the individuals in the sample, the deceased are not excluded from the individuals that have to pay. If $x\%$ is the share of surviving individuals, we only allocate $x\%$ of the total costs to them. Our approach boils down to the assumption that the share of the costs that we allocate to the dead ($(1-x)\%$) would in reality have been paid by young newcomers in our sample, that we did not include and that we suppose to be approximately equal to the number of dead.

with $\sum_p \pi_{ip} = 1$ for all i.

In principle we are interested in the well-being of individuals over the whole of their lives. The baseline approach in eq. (16) disregards the fact that equivalent incomes are correlated over time. Since we have no information about intertemporal preferences, we cannot aggregate over time at the individual level in a sophisticated way. A simple sensitivity analysis is possible, however, in which we assume that there is no time preference and calculate life-cycle well-being as the sum of the equivalent incomes obtained by the individual during the ten years. This yields

$$SW^{ex\,post} = \frac{1}{1-\rho} \sum_p \sum_i \pi_{ip} (\sum_t y^*_{ipt})^{1-\rho}. \tag{17}$$

4 Results

4.1 Basic results

Table 3 shows that the placebo strategy A leads to the poorest health results over the whole period and, importantly, that strategy C is better than strategy B in the second half of the period. This is in line with the findings of HAS.[20] The second panel of Table 3 gives the equivalized incomes, averaged over the surviving individuals only. Strategy C leads to a larger total cost of treatment, hospitalization and follow-ups. and this is reflected in lower individual incomes. The placebo strategy A is cheaper in the beginning of the period but leads to lower incomes at the end of the period, due to the increase in treatment and hospitalisation costs from the more frequent occurrence of cardiovascular events.

The resulting average *equivalent* incomes are shown in the bottom panel of Table 3. These results can be interpreted as the evaluation of the social welfare functions (16) and (17)

[20]The steady decline of SAH over time for the three strategies (table 3) is explained by the ageing of our sample, i.e. by the fact that we do not take up young newcomers. Moreover, individuals that die remain in the sample with a value of SAH equal to zero.

	t=1	t=2	t=3	t=4	t=5	t=6	t=7	t=8	t=9	t=10
	SELF-ASSESSED HEALTH OVER TIME									
	Individuals with hypertension									
A	65.053	63.437	61.801	60.083	58.346	56.586	54.788	53.002	51.254	49.524
B	65.148	63.645	62.120	60.523	58.903	57.251	55.562	53.875	52.198	50.527
C	65.142	63.637	62.112	60.516	58.900	57.252	55.571	53.894	52.223	50.561
	All individuals									
A	71.279	70.191	69.090	67.954	66.791	65.609	64.408	63.206	62.011	60.813
B	71.309	70.256	69.190	68.091	66.964	65.817	64.650	63.479	62.306	61.126
C	71.307	70.254	69.187	68.089	66.963	65.817	64.653	63.485	62.314	61.137
	AVERAGE EQUIVALIZED INCOMES OVER TIME									
	Individuals with hypertension									
A	1325.78	1327.20	1328.25	1329.42	1330.72	1331.95	1332.98	1334.77	1334.85	1334.02
B	1322.58	1324.53	1326.22	1328.12	1330.01	1331.65	1333.13	1335.24	1335.50	1334.80
C	1321.55	1323.42	1325.02	1326.92	1328.84	1330.53	1332.12	1334.35	1334.73	1334.06
	All individuals									
A	1335.19	1331.45	1330.69	1331.30	1332.17	1333.20	1333.98	1335.56	1335.51	1334.59
B	1332.01	1328.76	1328.65	1329.99	1331.45	1332.89	1334.13	1336.03	1336.15	1335.37
C	1331.02	1327.67	1327.47	1328.79	1330.28	1331.78	1333.12	1335.14	1335.38	1334.64
	AVERAGE EQUIVALENT INCOMES OVER TIME									
	Individuals with hypertension									
A	1214.86	1188.78	1161.76	1133.11	1104.35	1075.19	1045.18	1015.47	984.28	952.81
B	1213.54	1189.70	1164.95	1138.74	1112.17	1084.90	1056.82	1028.74	998.62	968.03
C	1212.46	1188.48	1163.59	1137.35	1110.82	1083.63	1055.70	1027.83	997.90	967.41
	All individuals									
A	1260.80	1249.88	1236.36	1226.31	1213.52	1202.93	1186.73	1168.50	1149.75	1131.47
B	1258.36	1248.40	1235.96	1227.02	1215.27	1205.53	1190.22	1172.72	1154.42	1136.53
C	1257.46	1247.44	1234.95	1226.02	1214.33	1204.66	1189.45	1172.08	1153.87	1136.07

Table 3: Self-assessed health, Equivalised incomes (after substracting out-of-pocket costs and treatment costs) and Equivalent incomes over time

for $\rho = 0$. Indeed, with zero inequality aversion, social welfare becomes the sum of equivalent incomes. Since the population remains constant in our calculations, the sum and the average yield identical rankings. Because the number of deceased matters for the social evaluation of the strategies, we did include the dead in the calculation of the averages: they are ascribed zero equivalent income. The larger costs with strategy C weigh more heavily in the individual preferences than the resulting better health. Strategy B therefore yields a higher level of social welfare than strategy C, and both are better than the placebo strategy A, except in the first periods.

Further calculations (see the online appendix) show that the ranking of the strategies does not change with increasing inequality aversion. Strategy B is better than strategy C in all periods. Moreover, with also with the lifetime approach (see eq. (17)), strategy B is preferred over strategy C and both are preferred over placebo.

A direct perspective on poverty is obtained by considering the sum of the equivalent incomes of the 30% worst-off in the different scenarios. In the left panel of table 4, the dead are included (with equivalent incomes = 0): in this case, strategy B is the best in all periods (and strategy C is better than the placebo strategy from period 6 onwards). Strategy B is also the best if we use the sum of the equivalent incomes over the ten periods to identify the worst-off. Since we care about the number of deceased, these are the most relevant results. For illustrative purposes, we show in the right panel of the table the results when we exclude the dead, i.e. when we consider in each period (and for the sum of the ten periods) those 30% surviving individuals with the lowest equivalent incomes. Under this assumption, the placebo-condition (resulting in higher equivalized incomes) is the best strategy up until period 8.

The conclusion is clear: for all degrees of inequality aversion and regardless of whether we take a period-by-period or a life cycle approach, strategy B (the cheapest strategy) is to be preferred over strategy C (most effective in health terms), and both are preferred over placebo. As shown in the online appendix, a traditional CEA (taking SAH as a measure of QALY) yields the same ranking of the strategies. To illustrate the additional insights obtained by using our richer evaluative framework, the next section presents some simulation exercises.

Dead individuals are included				Dead individuals are excluded			
Period	A	B	C	Period	A	B	C
1	17 305	17 307	17 295	1	18 317	18 277	18 264
2	16 464	16 517	16 475	2	18 498	18 462	18 411
3	15 431	15 573	15 580	3	18 587	18 558	18 505
4	14 540	14 633	14 604	4	18 679	18 659	18 644
5	13 398	13 411	13 367	5	18 774	18 753	18 751
6	12 500	12 623	12 612	6	18 949	18 932	18 917
7	11 400	11 590	11 567	7	19 010	18 970	18 966
8	10 336	10 543	10 547	8	19 062	19 096	19 084
9	9 268	9 449	9 459	9	19 127	19 128	19 116
10	8 204	8 447	8 442	10	19 202	19 204	19 193
sum over 10 years	146 107	147 051	146 937	sum over 10 years	188 805	188 670	188 532

Table 4: Results for the 30% worst-off

4.2 Some simulation exercises

4.2.1 Ex ante versus ex post

The difference the ex post- and the ex ante-perspectives becomes salient as soon as we introduce inequality aversion in the social welfare function. The results for an ex ante-evaluation with $\rho = 3$ are shown in Table 5. For $\varepsilon = 0$, moving from an ex post- to an ex ante-perspective has a strong effect on the evaluation of the placebo-strategy A, which now becomes the preferred strategy up until period 7. The absence of antihypertensive treatment increases the number of individuals that experience a cardiovascular event ex post - this increased number of worst-off individuals is weighted heavily in an inequality averse ex post social welfare function. However, in an ex ante perspective, this is not interpreted as a larger number of sick individuals but as a larger probability of getting sick. The differences in the *expected* outcomes between placebo and treatment strategies will be less pronounced (the lowest expected health outcome will definitely be larger than the lowest actual health outcome), leading to a weaker effect of the inequality aversion in the social welfare function. In the first periods (with the lowest probabilities of cardiovascular events) the much smaller costs of strategy A, which are divided over the whole population, may therefore dominate. This effect becomes even stronger when $\varepsilon = 2$. Under that assumption, the placebo strategy becomes the best in all periods (and strategy C is now preferred over strategy B).

Using $\varepsilon = 0$				Using $\varepsilon = 2$			
Period	A	B	C	Period	A	B	C
1	-0.00373	-0.00383	-0.00386	1	-18 500 000	-19 000 000	-18 800 000
2	-0.00379	-0.00388	-0.00391	2	-6 188 390	-6 330 165	-6 290 433
3	-0.00402	-0.00409	-0.00413	3	-2 889 736	-2 958 383	-2 940 238
4	-0.00429	-0.00435	-0.00439	4	-1 470 065	-1 503 356	-1 495 035
5	-0.00463	-0.00466	-0.00470	5	-857 624	-877 893	-872 949
6	-0.00503	-0.00505	-0.00509	6	-533 121	-546 459	-543 332
7	-0.00556	-0.00554	-0.00558	7	-293 869	-301 815	-300 126
8	-0.00625	-0.00620	-0.00624	8	-190 308	-195 594	-194 540
9	-0.00720	-0.00711	-0.00716	9	-133 953	-137 714	-136 984
10	-0.00865	-0.00850	-0.00855	10	-98 730	-101 511	-101 006

Table 5: Ex ante evaluation (values of the social welfare function) for $\rho = 3$

4.2.2 The welfare effects of social insurance

In our base case distributional considerations have only a minor effect on the ranking of the different strategies. This is explained by the relatively small cost of the antihypertensive treatments and by the distribution of the total cost over all citizens roughly in proportion to their income. However, we can also simulate more drastic distributional policies. As an example, consider the hypothetical situation where individual patients are required to pay the full cost of their treatment themselves as long as their income remains positive. More specifically, if y_i is the income of individual i and c_i is his treatment cost, we assume that the patient has to pay $min(c_i, y_i)$ and society takes care of the remainder. These remaining costs are divided over the top two deciles of the income distribution, in proportion to their income. We interpret this situation as that of a *laissez-faire* society, where the rich show sufficient charity to pay for the health care expenses of the poor, but only after the latter have first exhausted their own income. Table 6 shows that if society does not care about ex post-inequality ($\rho = 0$), the *laissez-faire* is slightly better than the social insurance system. However, this finding is completely reversed for $\rho = 3$. For all strategies a large welfare gain is realised by introducing a social insurance system. As a matter of fact, whereas most of the health economic analysis (for good reasons) focuses on a comparison of strategies B and C, our results show that the distributional effects of introducing social insurance are of a much larger order of magnitude. Strategy B is better than strategy C, even if patients have to pay their own expenses, but strategy C in a

$\rho = 0$							
Social Insurance (baseline)				Patients pay their own expenses			
Period	A	B	C	Period	A	B	C
1	3 042 303	3 036 432	3 034 252	1	3 042 410	3 038 062	3 036 238
2	3 015 956	3 012 392	3 010 064	2	3 016 310	3 014 817	3 012 983
3	2 983 345	2 982 376	2 979 942	3	2 984 248	2 985 378	2 983 476
4	2 959 082	2 960 789	2 958 391	4	2 960 513	2 964 249	2 962 399
5	2 928 233	2 932 443	2 930 174	5	2 929 759	2 935 894	2 934 069
6	2 902 677	2 908 942	2 906 837	6	2 904 426	2 912 641	2 910 902
7	2 863 568	2 871 990	2 870 152	7	2 865 075	2 875 516	2 873 901
8	2 819 588	2 829 769	2 828 224	8	2 820 760	2 833 022	2 831 549
9	2 774 343	2 785 605	2 784 281	9	2 774 983	2 788 290	2 786 809
10	2 730 230	2 742 457	2 741 328	10	2 730 174	2 744 549	2 743 162
$\rho = 3$							
Social Insurance (baseline)				Patients pay their own expenses			
Period	A	B	C	Period	A	B	C
1	-1712.19	-1670.04	-1675.16	1	-1734.79	-1684.48	-1688.72
2	-3242.70	-3152.69	-3161.18	2	-3277.70	-3174.33	-3181.29
3	-4810.61	-4674.07	-4685.40	3	-4849.69	-4698.06	-4707.87
4	-6449.39	-6263.91	-6277.13	4	-6496.60	-6292.36	-6303.94
5	-8149.22	-7917.24	-7931.64	5	-8201.34	-7948.61	-7961.56
6	-9646.54	-9371.70	-9386.58	6	-9704.05	-9406.54	-9420.01
7	-11435.03	-11117.74	-11131.63	7	-11500.73	-11157.50	-11170.20
8	-13237.56	-12882.56	-12894.55	8	-13302.56	-12922.11	-12933.45
9	-15053.92	-14672.19	-14682.71	9	-15117.19	-14711.08	-14722.03
10	-16892.07	-16488.16	-16496.63	10	-16961.03	-16530.91	-16540.02

Table 6: Welfare effect of social insurance

system of social insurance is much better than strategy B in the *laissez-faire*. Introducing distributional considerations through a broader notion of well-being points therefore to the crucial importance of the broader institutional setting.

5 Conclusion

In this paper, we have shown that it is possible to implement a richer approach than CEA or traditional CBA, in which individual preferences are respected without becoming subjectively welfarist (by using the notion of equivalent income) and in which an adequate specification of the social welfare function is used to tackle fairness issues. Our paper therefore counters an argument which is sometimes offered in defence of CEA, i.e. that

it is unrealistically ambitious to combine the detailed and complex medical information needed for a careful analysis of different treatments into a rich (and therefore also complicated) evaluative framework. Our evaluation of alternative antihypertensive treatments in France shows that it is possible to combine detailed medical information with survey results on the preferences of citizens concerning health and income. The former type of information is application-specific as is also the case in traditional CEA. The latter type of information, however, is generic. It is not necessary to organise a new survey of the population for each specific application. Pragmatic arguments in favour of ethically poorer approaches are not justified. That is the main message of this paper.

Working within a richer evaluative framework offers scope for a true ethical debate on crucial assumptions. Obvious examples are the choice of a specific value for the parameter of inequality aversion and the choice between the ex ante and the ex post-approach, where the latter takes into account the consequences of bad luck in the measurement of inequality. As illustrated by our application, sensitivity analysis can add useful information to this ethical debate.

Since implementation of our distribution-sensitive evaluative framework is not an unrealistic goal, it is useful to work further in order to improve the methodology for future applications. The main challenge is a better estimation of individual preferences concerning income, health and (ideally) other dimensions of life. We asked the retrospective willingness-to-pay to be in perfect health for the last 12 months. It is possible that, to a limited degree, respondents include in their responses on SAH and willingness-to-pay the long-run effects of the illnesses they suffered from in the past 12 months, including their effect on life expectancy (and the anxiety about life expectancy). Yet it is likely that our survey has yielded a severe underestimate of the true willingness-to-pay to be in perfect health for a health problem like hypertension, for which the long-run effects are essential. This may explain why the cost effects are larger than the health effects in our application. Of course, the standard CEA-approach also relies on techniques such as the standard gamble or time trade-offs to measure the subjective value of health. These techniques are not immune to "biases" either. Improving the estimation of individual preferences should be one of the first priorities on the research agenda.

References

Blackorby, C. and Donaldson, D. (1990) A review article: the case against the use of the sum of compensating variations in cost-benefit analysis. Canadian Journal of Economics 23, 471-494.

Brekke, K. (1997) The numéraire matters in cost-benefit analysis. Journal of Public Economics 64, 117-123.

Caussat L., Le Minez S., and Raynaud D. (2005) L'assurance maladie contribue-t-elle à redistribuer les revenus? Paris: DREES.

Decancq, K., Fleurbaey, M. and Schokkaert, E. (2015a) Inequality, income and well-being. In: Atkinson, A. and Bourguignon, F., (Eds.) Handbook of Income Distribution, Vol. 2A, pp. 67-140. New York: Elsevier.

Dreze, J.P. (1998) Distribution matters in cost-benefit analysis: comment. Journal of Public Economics 70, 485-488.

Fleurbaey, M. (2008) Fairness, responsibility and welfare, Oxford: Oxford University Press.

Fleurbaey, M. and Blanchet, D. (2013) Beyond GDP. Measuring welfare and assessing sustainability, Oxford: Oxford University Press.

Fleurbaey, M., Luchini, S., Muller, C. and Schokkaert, E. (2013) Equivalent income and the economic evaluation of health care. Health Economics 22, 711-729.

Fleurbaey, M., Luchini, S., Schokkaert, E. and Van de Voorde, C. (2012) Evaluation des politiques de santé: pour une prise en compte équitable des intérets des populations. Economie et Statistique No 455-456, 11-36.

Fleurbaey, M. and Zuber, S. (2015) Fair management of social risk. Mimeo.

Godet-Thobie, H., Vernay, M., Noukpoape, A. et al. (2008) Niveau tensionel moyen et prévalence de l'hypertension artérielle chez les adultes de 18 à 74 ans. Bulletin Epidémiologique Hebdomadaire 49-50, 478-483.

HAS, Haute Autorité de Santé (2012). Évaluation par classe des médicaments antihypertenseurs. Saint-Denis La Plaine: HAS.

Hausman, D. (2007) What's wrong with health inequalities? Journal of Political Philosophy 15, 46-66.

Johansson, P.-O. (1998) Does the choice of numéraire matter in cost-benefit analysis? Journal of Public Economics 70, 489-493.

Loewenstein, G. and Ubel, P. (2008) Hedonic adaptation and the role of decision and experienced utility in public policy. Journal of Public Economics 92, 1795-1810.

Schokkaert, E., Van de Voorde, C., Dormont, B., Fleurbaey, M., Luchini, S., Samson, A.-L. and Thébaut, C. (2013) Equity in health and equivalent incomes. In: Rosa Dias, P. and O'Donnell, O., (Eds.) Health and Inequality (Research on Economic Inequality, Volume 21), pp. 131-156. Bingley: Emerald.

Sen, A. (1985) Commodities and capabilities, Amsterdam: North-Holland.

Wagner, A., Sadoun, A., Dallongeville, J., Ferrieres, J., Amouyel, P., Ruidavets, JB., et al. (2011) High blood pressure prevalence and control in a middle-aged French population and their associated factors: the MONA LISA study. Journal of Hypertension 29, 43-50.

Fairness in cost-benefit analysis: a methodology for health technology assessment
Online Appendix - Not for publication

26th July 2016

Anne-Laure Samson, Erik Schokkaert, Clémence Thébaut, Brigitte Dormont, Marc Fleurbaey, Stéphane Luchini, Carine Van de Voorde[1]

[1]ALS: Université Paris-Dauphine; ES: Department of Economics, University of Leuven and CORE, Université catholique de Louvain; CT: Haute Autorité Santé, France; BD: Université Paris-Dauphine; MF: Princeton University; SL: Aix-Marseille University (Aix-Marseille School of Economics), CNRS and EHESS; CVDV: Department of Economics, University of Leuven.

1 Description of the survey

The survey contained the usual questions on demographic and socioeconomic characteristics of the individual and his/her household (gender, age, marital status, level of education, profession, level of monthly household income before taxes). We measure equivalized income y_i as the reported household income, divided by the modified OECD-scale (with weights 1 for a single, 0.5 for each additional adult in the household and 0.3 for each child below the age of 14). It is crucial to avoid confusion between this "OECD-equivalized income" and the "equivalent income", that was introduced in section 2 of the main text.

The questionnaire continued with detailed questions on specific diseases and health problems that the respondent might have experienced in the previous twelve months. Each respondent was presented a detailed list of 45 diseases grouped in 15 categories (e.g., respiratory diseases, cardio-vascular diseases, ...) and was asked whether he was affected by this disease during the last 12 months and whether he had been prescribed a treatment or not.[2] Open-ended questions were added for each of the 15 groups of diseases, in order to identify if the respondent had suffered from another disease, not specifically mentioned in the list. At the end of that section came a question on overall self-assessed health (SAH), where use was made of a visual 0-100 scale:

> "In the previous questions, you have indicated the health problems you have suffered from during the last twelve months. Taking this into account, can you now evaluate your health during the last twelve months on a scale from 0 to 100 (where 100 is the best possible health state and 0 is death)."

The SAH question was followed by a series of questions on health care use in the past twelve months, including non-reimbursed out-of-pocket payments. We also asked if the respondent has the "ALD" (chronic disease) status and if so, for which disease,[3] and if he

[2]The list of diseases was taken from the Enquête Santé et Protection Sociale (Health, Health Care and Insurance Survey) of IRDES (Institute for Research and Information in Health Economics).

[3]In the French system, patients who suffer from a disease which is classified as "chronic" are fully covered by the national health insurance for all the health care related to this disease. Examples are diabetes, heart failure, stroke or Alzheimer.

benefits from complementary insurance coverage (through employers or as an individual voluntarily purchased insurance) or from the "universal medical coverage" (Couverture Maladie Universelle Complémentaire CMU-C, which provides a free complementary coverage for individuals with low incomes). Finally, individuals were asked about their lifestyles (smoking habits, alcohol consumption, weight and height, ...)

After respondents had in this way considered their own economic and health situation, they were confronted with a retrospective willingness-to-pay question (meant to measure $WTP_i(h_i \to h^*)$). Interviewers first introduced a hypothetical scenario:

> "Imagine now that you would not have had any health problem during the last twelve months. In that case you would have been in perfect health and your quality of life would have improved. (We talk here only about the last twelve months and not about the potential improvement of your future health). Compared to your actual life experience during the last twelve months, would you have preferred not to have had the health problems that you had but with a reduction of your income (on top of the €x that you now have paid already as non-reimbursed care expenditures)."

Respondents could answer "yes", "no" or "do not know". Those who answered "no" were asked for further information about their reasons. Some of them were ready to answer "yes" after some additional explanation. All those who answered "yes", were then asked:[4]

> "What is the maximal amount of monthly income you would have accepted to give up under these conditions (i.e. in exchange for being in a state of perfect health during the last twelve months)?"

As a help, the respondents were shown payment cards (ranging from "less than €15" to "more than €1500"). After having been shown the cards, they were asked an open question about the exact amount of income they were willing to give up.

[4]It was made clear in the questionnaire that the relevant income concept was the equivalized income, i.e. the monetary income adjusted for household size.

We analysed the reasons given by respondents for answering "no" to the willingness-to-pay question. Respondents who answered that they did not want to give up any income because "my living standard is already so low that I cannot imagine to have less, even with perfect health" or that answered "other aspects of my life are more important for me than my health" are included in our analysis as having a true willingness-to-pay equal to zero. Protest voters are those who answered that the question was too difficult or who, even after further explanation, kept to the position "it is not my duty to pay for a better health". They were removed from the analysis. Because the protest voters differed from the rest of the sample in terms of observable variables (e.g., there is a higher proportion of females), we introduced a selection equation in the estimation of the preferences. However, for the most crucial variables the selection bias is small. There are no significant differences for the prevalence of hypertension. After the removal of the incomplete and protest answers, the sample used in our analysis consists of 2413 individuals.[5]

2 Estimation of preferences[6]

With the information that is available in the survey, we can now directly implement eq. (5) to compute equivalent incomes for all individuals in the sample. To simulate the effect of different policies on equivalent incomes, we need in addition information about preferences. Estimating preferences at the individual level is impossible with our data, however, since the only information available for each individual is that his/her actual situation (y_i, h_i) and the hypothetical situation $(y_i - WTP_i, h^*)$ are on the same indifference curve. To make progress we have to combine information obtained from different individuals and make the assumption that preferences are homogeneous at the group level. In order not to push the data into a straitjacket, we opted for a flexible functional form, allowing for coefficients that differ according to age and gender. We measure income as the OECD-equivalized income and health as self-assessed health (SAH).

[5] We removed 13 individuals that were either 18 years or more than 96 years old, because for these age groups the HAS-model does not contain probabilities of getting hypertension or incurring a cardiovascular event.

[6] More information on the estimation procedure, including more detailed estimation results, can be found in Schokkaert et al. (2013).

More specifically, we specify the willingness-to-pay of individual i to be in perfect health as

$$WTP_i = \alpha_i(1-h_i) + \beta_i(1-h_i)^2 + \gamma_i y_i(1-h_i) + \delta_i y_i^2(1-h_i) + \mu_i y_i(1-h_i)^2 + \varepsilon_i \quad (1)$$

where ε_i is an idiosyncratic disturbance term with mean zero. Interindividual differences in the marginal rate of substitution between income and health are modelled by varying the coefficients in eq. (1):

$$\begin{aligned} \alpha_i &= \alpha_0 + \alpha_A age_i + \alpha_M male_i \\ \beta_i &= \beta_0 + \beta_A age_i + \beta_M male_i \\ \gamma_i &= \gamma_0 + \gamma_A age_i + \gamma_M male_i \\ \delta_i &= \delta_0 + \delta_A age_i + \delta_M male_i \\ \mu_i &= \mu_0 + \mu_A age_i + \mu_M male_i \end{aligned} \quad (2)$$

where age_i refers to the age of individual i and $male_i$ is a dummy, taking the value 1 for males. The functional form (1) does impose neither monotonicity nor quasi-concavity of the utility function. However, it does impose that the expected willingness-to-pay for an individual in perfect health (i.e. with $h_i = h^* = 1$) is equal to zero. This theoretical constraint makes our specification consistent with the theoretical framework sketched in section 2 of the main text.

As announced in the presentation of the data, we estimated the parameters of eqs. (1)-(2) with a two-step procedure to take into account that the protest voters are a selected sample.[7] The first step is a probit selection equation for refusing to give a WTP answer as a function of health, income, age and gender. In the second step we estimated eqs. (1)-(2) with OLS including the inverse Mill's ratio derived from the probit equation as an additional regressor. Of course, this whole procedure is just a curve-fitting exercise and should not be seen as the testing of any theory. While an F-test shows that the overall fit of equation (1) is satisfactory ($p < 0.0001$), it is impossible to interpret the individual coef-

[7]For this estimation, we removed 98 observations with SAH < 20 or with an income > €4000. Not surprisingly, our flexible functional form gives strange results for this range of the variables, where we have very few observations.

ficients in a meaningful way given the highly nonlinear specification of eq. (1) and the large degree of multicollinearity between the right-hand side variables. It is more informative to look at the indifference curves that can be computed on the basis of the estimated parameters, taking into account that utility is given by $y_i^*(y_i, h_i) = y_i - WTP_i(y_i, h_i \rightarrow h^*)$.[8] These indifference curves are shown in Figure 1 for the three age quartiles. Overall, they look reasonably good, certainly for income levels above €1000 (the median income in the full sample is €1200). They are rather flat, however, especially at very low incomes. Of course, this just reflects the finding that the WTP-answers in the survey are rather low.

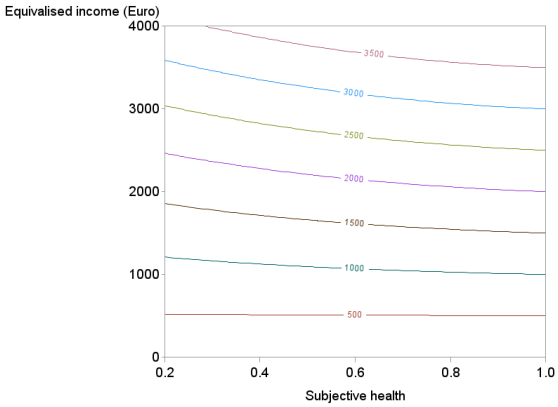
3 Empirical procedure: evaluation of the three strategies

We now describe the additional empirical steps that are needed for the assessment of the three strategies that were described in section 3-1 of the main text. We choose to evaluate the strategies on a horizon of ten years. From now on, we will therefore introduce in our notation a subscript τ to indicate the period considered, with $\tau = 0$ standing for the situation at the moment of our survey. In section 3.1 we show how we used the HAS prediction model to define the possible health trajectories that the individual can experience during the period of ten years and to associate with each of these trajectories individual-specific probabilities of occurrence. We then explain how we calculate equivalent incomes. Each individual starts in period 0 with his observed income and SAH. Depending on the followed trajectories we adjust dynamically his health (section 3.2) and income levels (3.3) over time. This enables us to compute, for each individual and period, the equivalent income for every possible trajectory. We can then compute the ex post social welfare for each period for the three strategies A, B and C. In section **??** we discuss how we introduced the time dimension in the calculation of social welfare.

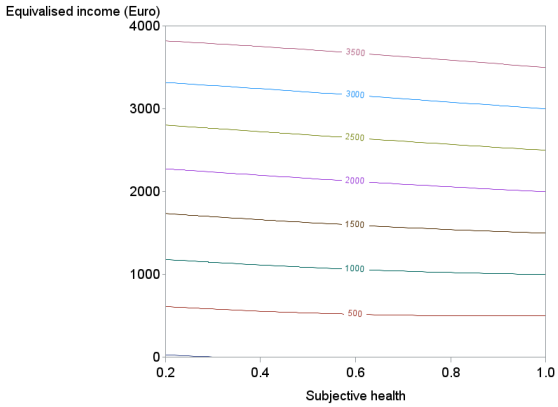
3.1 The distribution of events with the different strategies

Individuals with hypertension can over the considered period of ten years experience different cardiovascular events (stroke, angina, myocardial infarctus, heart failure) as well as

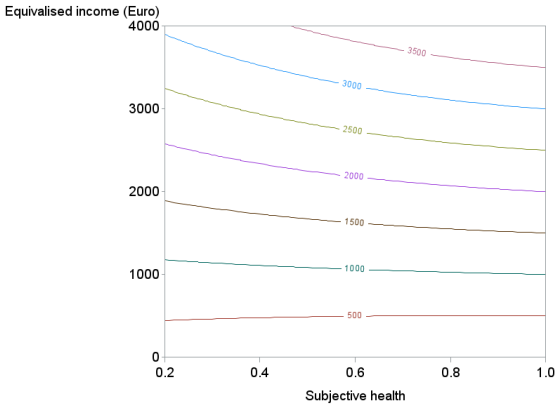
[8] The estimates of the individual coefficients are reported in Schokkaert et al. (2013).



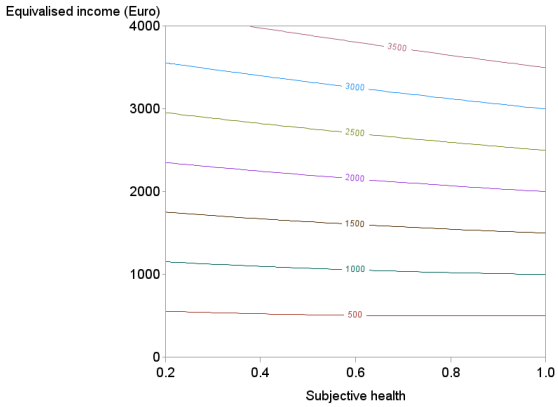
Males, age 39



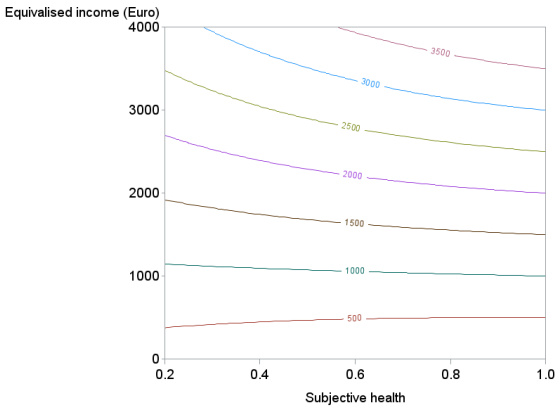
Females, age 39



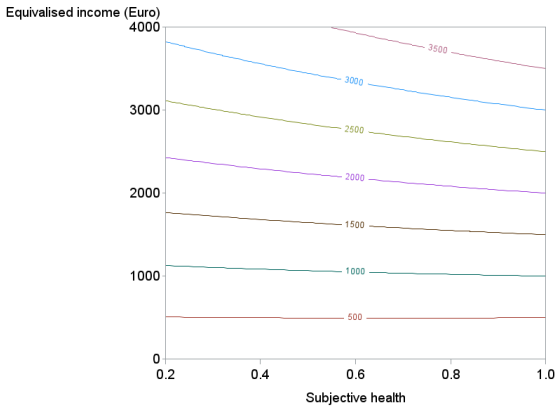
Males, age 54



Females, age 54



Males, age 66



Females, age 66

Figure 1: Estimated indifference curves

renal failure and end-stage renal failure. They can die as a consequence of the event they experience, or they can survive. Over the whole time span, i.e. 10 years, an individual can experience 3376 different trajectories.[9] The trajectories are illustrated for the first two periods in the decision tree of Figure 2.[10]

We calculated the probabilities of all these trajectories, using the information present in the HAS model. These probabilities obviously depend on the chosen treatment strategy. The HAS model works with two kinds of probabilities. First, there are the probabilities of experiencing a first cardiovascular or renal failure event. This probability of experiencing a first event is increasing with time (whatever the event). With the placebo strategy the risks depend on the individual's characteristics: gender, age, diabetes and smoking habits. In the case of antihypertensive treatment, these risks do no longer depend on individual characteristics, but on the prescribed line of treatment (the probability of controlling blood pressure with tritherapy being fixed at 1). Second, there are the probabilities of experiencing a second event in period τ, given that a first event occurred in period $\tau - h$, with $h \geq 1$. The probability of occurrence of a second event depends on the occurrence of a first event in a previous year and on the nature of this first event (stroke, angina, myocardial infarctus, etc.).[11]

Aside from these probabilities of events, all individuals in the sample, with or without hypertension, may die from other causes. We implement an "all causes mortality rate" dependent on age and gender.[12]

[9] We cannot generate in our ex post approach a matrix that considers all possible trajectories and then give a probability of zero to those that are not relevant. Given the large numbers of possible trajectories, this would not be tractable. To tackle this problem, we worked with recursive functions and imposed restrictions for some trajectories that were not possible.

[10] The decision tree is simplified, because we do not show the arms "death from other causes" and "alive with no event".

[11] We simplify the HAS assumptions with respect to mortality following heart failure. Contrary to stroke and myocardial infarcts, which are acute events, heart failure is a chronic disease. In the HAS model individuals may therefore die from heart failure each year following the event. We use cumulative probabilities of dying for heart failure and apply them at the time of occurrence of the heart failure. This simplifying assumption does not change the total number of people who die, but has as a consequence that people die earlier.

[12] We used the mortality rates that were produced for 2009 by the French Institute for Demographic Studies (INED).

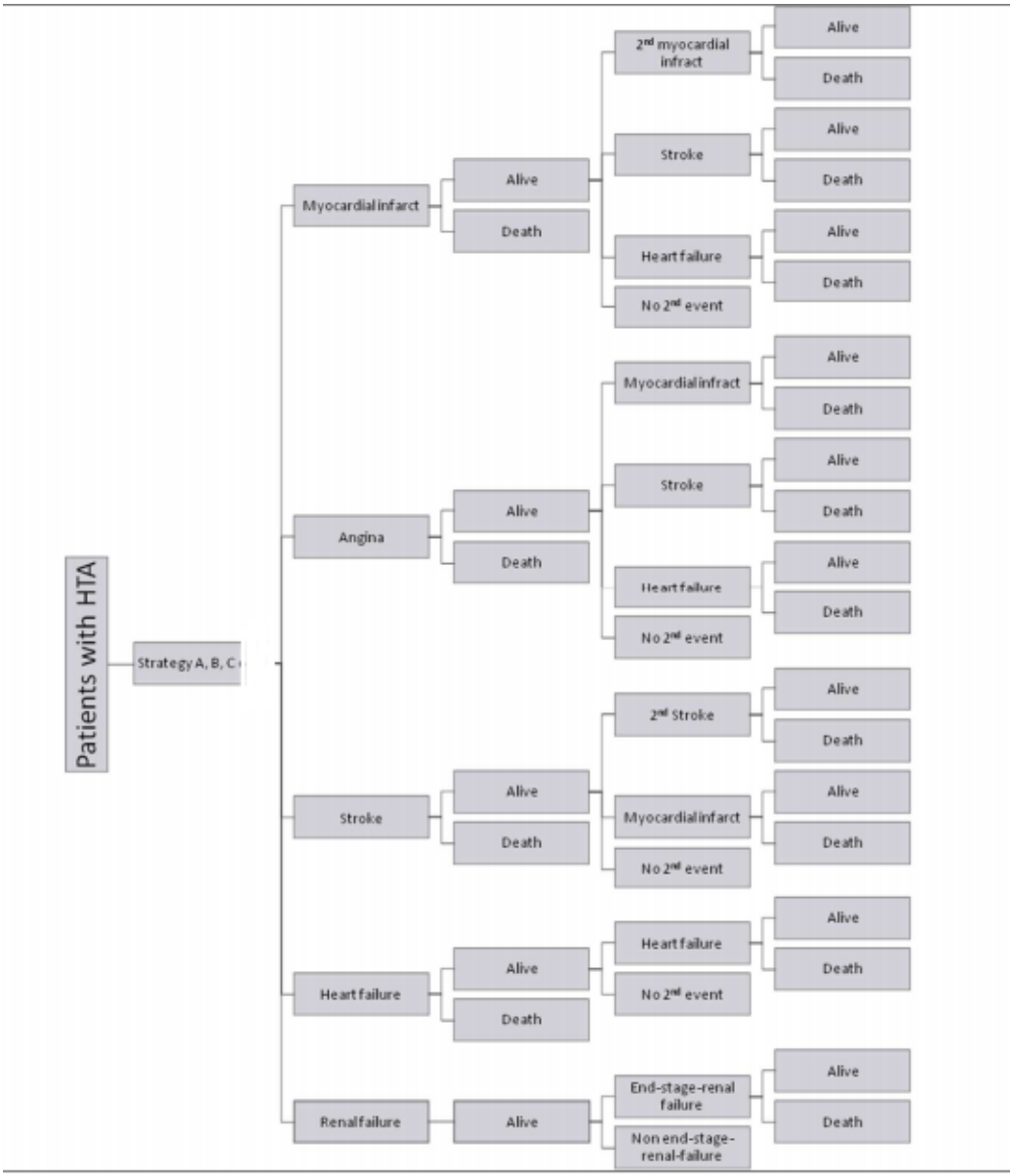


Figure 2: Decision tree

3.2 Impact of events on health: estimation results

The following linear regression is estimated on our sample:

$$SAH_{i,0} = c + \sum_j s_j ev_{ij,0} + a_1 age_{i,0} + a_2 age_{i,0}^2 + \sum_k b_k x_{ik,0} + u_{i,0}, \tag{3}$$

where the subscript 0 is introduced to indicate that the regression is run on our sample data (i.e. each individual is in period 0 of his 10-years trajectory), $ev_{ij,0} = 1$ if individual i experienced event j in the 12 months before the interview, $age_{i,0}$ is the age of individual i at the time of the interview, $x_{ik,0}$ indicates a list of control variables and $u_{i,0}$ is a disturbance term. The estimation results are shown in Table 1.[13] The coefficients of the control variables are in line with theoretical expectations and SAH decreases with age except for the very old (where we have only a limited number of observations and where there is also a selection effect at work). Myocardial infarctus, stroke and angina have a strong negative effect on SAH. In line with the literature, the effect of infarctus is larger than that for angina. Note that the other three events (renal failure, end-stage renal failure and heart failure) were not included in the list of 45 diseases in the questionnaire and could therefore not be included in eq. (3). On the basis of the information on the utility levels associated with all cardiovascular events, as reviewed in the technology assessment reports produced by NICE, we use the coefficient obtained for myocardial infarctus as a proxy for the impact of heart failure and renal disease, and the coefficient obtained for stroke as a proxy for the impact of end-stage renal failure.[14]

3.3 Impact of events on income: estimation results

We assume that health events do not influence the transfer payments (and hence the equivalized income) of non-working individuals. For the subsample of working individuals

[13] For this estimation, we did not exclude the protest voters from the sample.

[14] These are probably underestimates of the true effect. According to the National Clinical Guideline Centre (2011, p. 520), the utility level associated with heart failure (0.71) is lower than the utility level associated with myocardial infarctus (0.76 for the first six months), and the utility level associated with end-stage renal failure (0.60) is lower than the utility level associated with stroke (0.63).

		Coeff.	(Std Error)
Events	Angina	-10.469***	(2.068)
	Myocardial infarctus	-14.857***	(1.957)
	Stroke	-7.795***	(2.494)
Socio-demographic	Age	-0.643***	(0.102)
	Age Squared	0.0044***	(0.001)
	Male	4.523***	(0.679)
Education	No Diploma	-5.272***	(1.279)
	Primary School Certificate	-3.862***	(1.351)
	GCSE	-2.466**	(0.981)
	Baccalauréat	Ref.	Ref.
	University ($\leq$ 2 years)	1.370	(1.244)
	University ($\geq$ 3 years)	-1.098	(1.227)
	Other Diploma	-17.151***	5.496
Health Insurance	National Health Ins. only	-2.415*	(1.363)
	CMUC only	-2.366*	(1.400)
	Complementary Insurance	Ref.	Ref.
Family Situation	Marital Life	3.033***	(0.709)
	Single	Ref.	Ref.
	At least one child	1.001	(0.843)
	No child	Ref.	Ref.
Lifestyles	Smoker	-5.538***	(0.726)
	Underweight	-1.013	(2.118)
	Normal weight	Ref.	Ref.
	Overweight	-2.523***	(0.786)
	Obese	-8.514***	(0.939)
	Severely Obese	-14.454***	(1.676)
	No alcohol	Ref.	Ref.
	Alcohol - no risk	3.105***	(0.732)
	Alcohol - risky behaviour	-0.543	(1.610)
	Constant	92.968***	(2.439)
	R-squared	0.193	
	Nb of Obs.	3,304	

Table 1: Estimation results for SAH

		Coeff.	(Std Error)
Health	SAH	2.844***	(1.059)
Socio-demographic	Age	17.350***	(1.814)
	Male	426.911***	(39.180)
Education	No Diploma	-278.334***	(82.025)
	Primary School Certificate	-219.759*	(113.863)
	GCSE	-91.558*	(53.219)
	Baccalauréat	Ref.	Ref.
	University ($\leq$ 2 years)	-36.669	(63.391)
	University ($\geq$ 3 years)	216.039***	(70.340)
	Other Diploma	-304.834	496.085
Profession	Farmer	-626.104***	(165.735)
	Artisans / self-employed	386.424***	(77.739)
	Employee	Ref.	Ref.
	Top executive	683.138***	81.865
	Middle class profession	285.259***	(58.441)
	Workmen	-155.497***	(49.136)
Family Situation	Marital Life	66.508	(41.082)
	Single	Ref.	Ref.
	At least one child	31.063	(40.647)
	No child	Ref.	Ref.
	Live in Paris area	174.169***	(54.037)
	Constant	253.046**	(128.207)
	R-squared	0.361	
	Nb of Obs.	1,252	

Table 2: Estimation results for equivalized income

we estimate the following regression

$$y_{i,0} = d + hSAH_{i,0} + l.age_{i,0} + \sum_k m_k z_{ik,0} + v_{i,0}, \tag{4}$$

with $y_{i,0}$ indicating the level of monthly personal income, $z_{ik,0}$ the control variables and $v_{i,0}$ a disturbance term. The estimation results are shown in Table 2. An increase of 10 points for self-assessed health is associated with an increase in income of €28 (remember that average personal income in the sample is 1247€).

Different categories of individuals were now treated as follows:

1. For working individuals that are less than 65 years old, i.e. before the age of retire-

ment, the income in period τ is calculated as

$$y_{i,\tau} = y_{i,\tau-1} + \hat{h}(SAH_{i,\tau} - SAH_{i,\tau-1}) + \hat{l}. \quad (5)$$

The coefficient $\hat{l}$ captures the appreciation of income with age. The value of $SAH_{i,\tau}$ is first simulated using eq. (12) in the main text and then introduced in eq. (5).

2. Individuals that receive a transfer income in period 0 (pensioners, unemployed, housewives) keep the same level of transfer income during their whole trajectory. Individuals that are unemployed in the sample, are assumed to remain unemployed during the whole period of 10 years. Conversely, individuals that are employed in period 0 do not become unemployed.

3. Individuals that are employed in the sample but reach the age of 66 in period τ, get for that period a retirement pension equal to the average pension of all the individuals in the sample with the same profession and the same sex. For the periods $t = \tau + 1, \ldots, 10$, they are treated in the same way as the other pensioners.

4. Individuals who are defined as "students" in the sample are assumed to enter the labour market at an age corresponding to their type of studies. In that period they receive an income equal to the average income of the working individuals in the sample who are between 30 and 40 years old and have the same level of education and the same sex. Afterwards their income is adjusted with eq. (5).

5. For individuals that die in period τ, we assume that $y_{it} = 0$ for $t = \tau, \ldots, 10$.

4 Results: additional figures & tables

All the results in the paper are derived under the assumption that individuals that die are kept in the sample with an equivalent income equal to zero. It could be hypothesized that this underestimates the value of longevity. This is especially worrying in our setting, since concern about the long-run consequences of disease and the resulting anxiety about dying

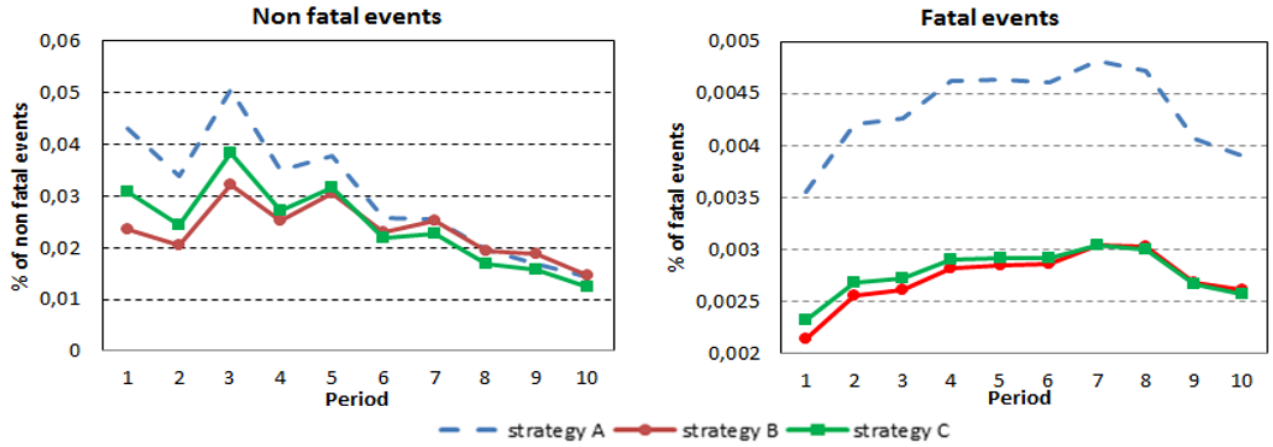


Figure 3: Percentage of non-fatal and fatal events

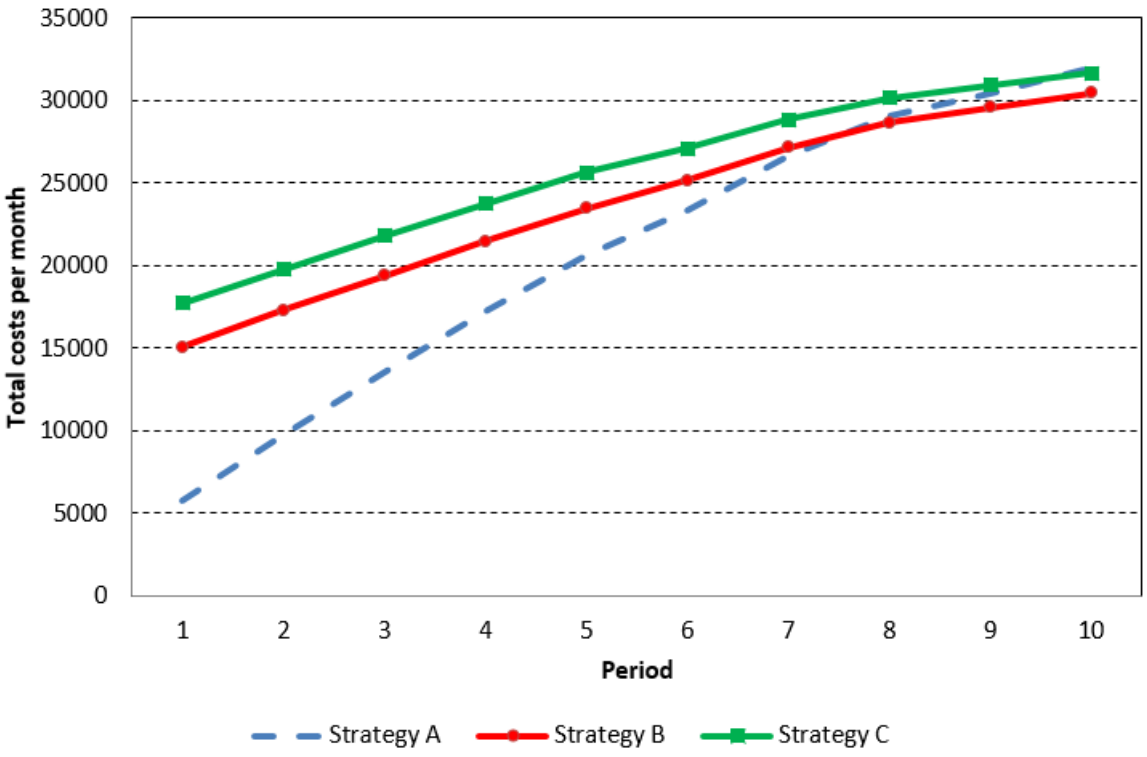


Figure 4: Treatment, hospitalization, follow-up costs (net of out-of-pocket payments)

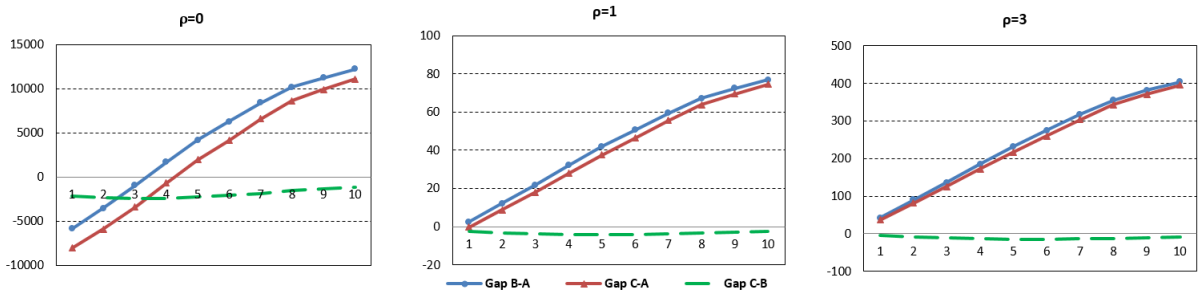


Figure 5: Differences in social outcomes for different values of ρ

	$\rho = 0$	$\rho = 1$	$\rho = 2$	$\rho = 3$
A	24 393 338 880	17 010 086	- 5 861 306	- 29 305 498
B	24 740 663 296	17 327 042	- 5 589 048	- 27 944 194
C	24 717 901 824	17 318 696	- 5 595 014	- 27 974 022

Table 3: Life cycle well-being - robustness check for longevity

are probably not well captured in our survey of preferences. We therefore performed a robustness check within the life-cycle approach, in which we give all individuals that die during the 10-year period an overall equivalent income of 0.1 (rather than the sum of the equivalent incomes during the years in which they lived). Table 3 shows that this does not change the ranking of the strategies.

5 Simulation exercises: detailed results

5.1 Ex ante versus ex post

Table 4 presents the results of the ex-ante evaluation when $\rho = 0$. The right part shows the results produced if we increase the degree of risk aversion ε to 2. Differences between the ex-post and the ex-ante perspectives are found when we introduce inequality aversion into the social welfare function. The results for $\rho = 3$ are discussed in the main text of the paper.

Using $\varepsilon = 0$				Using $\varepsilon = 2$			
Period	A	B	C	Period	A	B	C
1	3 043 379	3 039 263	3 037 576	1	-294.65	-286.23	-287.26
2	3 018 448	3 016 819	3 015 133	2	-600.69	-582.69	-584.39
3	2 987 154	2 987 921	2 986 182	3	-914.23	-886.93	-889.20
4	2 964 329	2 967 367	2 965 681	4	-1241.94	-1204.85	-1207.50
5	2 934 495	2 939 601	2 937 992	5	-1581.87	-1535.48	-1538.36
6	2 910 105	2 916 957	2 915 478	6	-1931.34	-1876.38	-1879.36
7	2 872 039	2 880 688	2 879 402	7	-2288.92	-2225.47	-2228.25
8	2 828 801	2 838 940	2 837 871	8	-2649.32	-2578.32	-2580.72
9	2 783 791	2 794 816	2 793 892	9	-3012.48	-2936.14	-2938.24
10	2 739 971	2 751 831	2 751 076	10	-3380.00	-3299.22	-3300.92

Table 4: Ex ante evaluation (values of the social welfare function) for $\rho = 0$

5.2 A cost-effectiveness analysis

The most popular criterion in cost-effectiveness analysis is the incremental cost-effective ness ratio (ICER) with the cost difference in the numerator and the health gains in the denominator. Our data allow us to mimic closely this approach. Since the placebo treatment A is dominated by strategies B and C we focus on the comparison of the latter two strategies and more specifically on the question whether the incremental health gains resulting from strategy C justify its larger cost. The traditional approach measures outcomes in term of the (unweighted) sum of QALY-gains. We do not have QALY's in our data, but for this illustration they can be reasonably well approximated by our measure of self-assessed health: it lies between 0 and 1 and we have a value for each year. There is a large literature on the pros and cons of discounting. Again, since this is only an illustration we will discount neither costs nor benefits. We can then approximate the ICER for C with respect to B as

$$ICER = \frac{\sum_{\tau=1}^{10}(TC_{\tau}^{C} - TC_{\tau}^{B})}{\sum_{\tau=1}^{10}\sum_{i}(SAH_{i,\tau}^{C} - SAH_{i,\tau}^{B})}. \tag{6}$$

The result is shown in the last column of Table 5. We find that strategy B is preferable over strategy C as long as one year in perfect health (i.e. with $SAH = 1$) is valued at less than €481,867. This value is much larger than the values that are commonly taken as the threshold in CEA: cost-effectiveness analysis therefore also suggests that it is not worthwhile to opt for the more expensive strategy C. Results year-by-year show that since

Period	1	2	3	4	5	6	7	8	9	10	**Global**
ICER	-	-	-	-	-	2 888.559	266.545	109.633	71.522	47.573	**481.867**

Table 5: Incremental cost-effectiveness ratios (health measured as SAH)

the differences in costs become smaller over time and the health gains larger, the period-by-period ICER for strategy C gets smaller over time. Extrapolating these results suggests that the overall ICER for strategy C would decrease if we used a longer time horizon.