

2011/43



Behavioral fair social choice

Marc Fleurbaey and Erik Schokkaert



CORE

DISCUSSION PAPER

Center for Operations Research
and Econometrics

Voie du Roman Pays, 34
B-1348 Louvain-la-Neuve
Belgium

<http://www.uclouvain.be/core>

CORE DISCUSSION PAPER
2011/43

Behavioral fair social choice

Marc FLEURBAEY¹ and Erik SCHOKKAERT²

August 2011

Abstract

Behavioral economics has shaken the view that individuals have well-defined, consistent and stable preferences. This raises a challenge for welfare economics, which takes as a key postulate that individual preferences should be respected. We agree with Bernheim (2009) and Bernheim and Rangel (2009) that behavioral economics is compatible with consistency of partial preferences and that subjective welfare measures do not offer an attractive way out of the impasse. However, their analysis which rests on traditional concepts like Pareto optimality and compensation tests, is not adequate to introduce distributional considerations. We explore how partial preferences can be introduced in the recent theory of welfare that has developed from the theory of fair allocation. We revisit the key results of that theory in a framework with partial preferences and show how one can derive partial orderings of individual and social situations.

Keywords: behavioral economics, preferences, welfare economics, psychology, social choice, fairness.

JEL Classification: D60, D71

¹ Princeton University, USA. E-mail: mfleurba@princeton.edu

² Department of Economics, Katholieke Universiteit Leuven, Belgium and Université catholique de Louvain, CORE, B-1348 Louvain-la-Neuve, Belgium. E-mail: erik.schokkaert@econ.kuleuven.be

This paper has benefited from detailed comments by François Maniquet, from discussions with D. Bernheim, A. Deaton, D. Kahneman, B. Tungodden, G. Weyl and from the reactions of audiences in Oxford, Maastricht and Leuven.

This paper presents research results of the Belgian Program on Interuniversity Poles of Attraction initiated by the Belgian State, Prime Minister's Office, Science Policy Programming. The scientific responsibility is assumed by the authors.

1 Introduction

One of the challenges for welfare economics is the formulation of adequate criteria to evaluate (re)distribution. Without such criteria policy evaluation can only be based on the Pareto criterion. Pareto-improving policy measures are rare, however. Rejecting all other policies leads to a conservative defence of the status quo, while the Kaldor-Hicks criterion of potential Pareto improvements is lacking ethical content. Indeed, the existence of a “potential improvement” is not very relevant if the necessary compensations remain purely hypothetical. To go beyond these Pareto-type approaches, one needs a concept of interpersonally comparable well-being. Traditional welfare economics has struggled for a long time with the issue of interpersonal comparisons. Arrow’s impossibility theorem has most often been interpreted as showing that the informational basis of ordinal preferences is insufficient to derive an ordering of social states. In the wake of Sen (1970), a large literature explored the consequences of going beyond such ordinal preference information and derived welfare criteria under different assumptions about interpersonal comparability and measurability of individual subjective welfare (for an overview, see d’Aspremont and Gevers, 2002). Yet choice behavior can only reveal ordinal preferences and, despite the growing popularity of the happiness literature, there remains much scepticism about the possibility to measure subjective welfare in a reliable way. Moreover, this so-called welfarist approach to social choice has been heavily attacked on ethical grounds (see, e.g., Dworkin, 1981; Sen, 1985). Importantly, as we will see, it does not respect individual preferences.

Recent developments in the theory of fair allocations have shown, however, that the common interpretation of Arrow’s theorem is wrong. It is possible to derive an interpersonally comparable concept of well-being based only on information about ordinal preferences – and to apply it for the analysis of specific policy measures (Fleurbaey and Maniquet, 2011). One attractive approach is based on the concept of equivalent income, which is firmly rooted in the tradition of money-metric utility (see also Fleurbaey, 2009). Although

this so-called fairness approach offers a promising way out of Arrow’s impossibility without necessitating the use of subjective utilities, it does rest on the assumption that well-defined individual preferences exist.

The findings of behavioral economics have cast doubt on this assumption. The existence of “behavioral anomalies” suggests that it is difficult to interpret individual choice behavior as the maximization of well-defined preferences. This has important implications for welfare economics. Some authors (Frey and Stutzer, 2002; Kahneman et al., 1997; Kahneman and Sugden, 2005; Köszegi and Rabin, 2008; Layard, 2005) have advocated to focus on experience utility (and subjective happiness) rather than on decision utility. This would bring us back to the welfarist interpretation of Arrow’s theorem. Other authors refuse to take this step and formulate preference-based welfare criteria (Bernheim and Rangel, 2009; Bernheim, 2009; Dalton and Ghosal, 2010; Rubinstein and Salant, 2009; Salant and Rubinstein, 2008). The proposed preference relations are incomplete and the question remains whether it is possible in this approach to go beyond the Pareto-criterion.

In this paper we bring together these two recent streams of literature. This allows us to raise two questions. First, suppose that one accepts the equivalent income approach, can it survive in a setting with incomplete preferences? Second, supposing that one works with an incomplete preference relation as defined, e.g., by Bernheim and Rangel (2009), is it then possible to derive an interpersonally comparable concept of well-being? We show that the answers to both questions are positive and complementary. Using the incomplete individual preference relation proposed by Bernheim and Rangel (2009), we derive an incomplete ordering of personal situations in terms of well-being and we argue that this concept of well-being, which relies only on ordinal preferences, can be used for distributional judgments. Respect for individual preferences is the key value in our approach and we explore how far one can go if one accepts this key value.

We briefly introduce in Section 2 the main findings about the equivalent income in a setting with well-defined and complete preferences. We then summarize in Section 3 some

relevant features of the behavioral welfare economic approach. In Section 4 we discuss the relevance of incomplete preference relations for describing individual's "authentic preferences" about what a good life is and we describe our own formal framework. Sections 5 and 6 extend the main results of the theory of fair social choice to the framework with incomplete preferences. Section 7 concludes.

2 Interpersonal comparisons and social rankings with well-defined preferences: the concept of equivalent income

We consider first the problem of ranking social states in a setting where individuals have well-defined ideas about what a good life is. The characteristics of a life (such as consumption, health status, employment status and so on) are summarized in a vector $x \in X$, and we assume for simplicity that $X = \mathbb{R}_+^\ell$. Each individual i has a (complete) preference ordering R_i over the vectors x_i , which reflects her informed judgment about what makes a life good or bad: $x_i R_i x'_i$ if i weakly prefers the life described by x_i to the life described by x'_i . Let $x_i P_i x'_i$ denote strict preference and $x_i I_i x'_i$ denote indifference. We assume that R_i is continuous and monotonic.¹ Let D be the set of preference orderings satisfying these assumptions.

We will proceed in two steps. First, we propose a method for interpersonal comparisons of individual well-being, relying only on ordinal preferences. Formally, this means that we look for a ranking of personal situations (x, R) in terms of well-being. Second, we consider a population with n individuals and we propose a method to derive social priorities, i.e., to devise a social ranking of allocations $(x_1, \dots, x_n)$. Again, we will rely only on information about ordinal preferences. The two problems are obviously linked: the interpersonally comparable concept of well-being, derived in the first step, will play an essential role in the second step.

¹The monotonicity assumption makes the exposition easier but the main results of this section can also be derived without it – see, e.g., Fleurbaey et al. (2009). Vector inequalities will be denoted $\geq, >, \gg$.

2.1 Interpersonal comparisons when preferences differ: equivalent income

Let us denote the ranking of personal situations (x, R) in terms of well-being by $\succsim$ (with asymmetric and symmetric parts $\succ, \sim$). We require this ranking to be transitive. Note that anonymity is assumed from the outset, so that the identity of individuals is not part of the description of situations. The statement $(x, R) \succsim (x', R')$ can be interpreted as stating that the well-being of an individual with preferences R in state x is at least as large as the well-being of an individual with preferences R' in state x' . Since we want to respect individual preferences, we impose the following Preference Principle as an essential requirement that interpersonal comparisons have to satisfy:

Preference Principle: $(x, R) \succeq (x', R)$ if xRx' .

This principle embodies the idea of individual sovereignty: if an individual prefers x to x' , then the well-being ranking should reflect this personal preference. It also embodies the idea of respecting interpersonal comparisons across individuals sharing the same preferences.² If two individuals have the same preferences R and agree that life situation x is better than life situation x' , the well-being ranking should assign to the individual in x a larger well-being level than to the individual in x' . It is important to see that this preference principle is incompatible with ranking the individual life situations on the basis of subjective well-being (or happiness): it is very well possible that two individuals agree that x is better than x' , and that at the same time the individual in situation x has a lower level of subjective happiness than the individual in situation x' , e.g., because she has more ambitious aspirations. Therefore ranking individual situations on the basis of happiness does not respect individual preferences.³

²In Fleurbaey et al. (2009), the two ideas are axiomatized separately as the “Personal-Preference” and the “Same-Preference” principles respectively.

³Fleurbaey et al. (2009) develop this basic insight and argue that happiness data can still be very useful to recover information about individual ordinal preferences.

The preference principle only bites if we have to compare the situations of two individuals with the same preferences. A more challenging problem of interpersonal comparisons arises when this is not the case. A natural starting point might seem to impose a dominance principle:

Dominance principle $(x, R) \succcurlyeq (x', R')$ if $x \geq x'$; $(x, R) \succ (x', R')$ if $x \gg x'$.

At first sight, it looks obvious that a situation x which is better than another situation x' in *all* dimensions should also lead to a larger level of well-being. However, the dominance principle does not bring us very far, since it is incompatible with the preference principle. Indeed, it implies that (x, R) is as good as (x, R') for all x and all R, R' , so that R plays no role in the evaluation of (x, R) . Even the weak requirement that (x, R) be better than (x', R') whenever $x \gg x'$, for our domain of monotonic preferences in all dimensions, is incompatible with the personal-preference principle. This is shown by the following example from Brun and Tungodden (2004). Assume $X = \mathbb{R}_+^2$ and take $x_i, x_j, x'_i, x'_j \in X$ and R_i, R_j such that $x_i \gg x_j, x'_i \ll x'_j, x'_i P_i x_i$, and $x_j P_j x'_j$. Figure 1 illustrates this configuration. The preference principle implies that (x'_i, R_i) is better than (x_i, R_i) and (x_j, R_j) is better than (x'_j, R_j) while the dominance principle implies that (x_i, R_i) is better than (x_j, R_j) and (x'_j, R_j) is better than (x'_i, R_i) . By transitivity, one obtains that (x_i, R_i) is better than (x_i, R_i) , which is impossible.

We are at a crossroads here. Either we keep the dominance principle and we relax (or give up) the preference principle.⁴ Or we keep the preference principle but then we cannot retain the dominance principle. In this paper we follow this second route. We can nonetheless seek a weakening of the dominance principle that is compatible with some respect for preferences. A natural approach consists in restricting the application of the dominance idea to a certain region B of the space X of life vectors x . As it turns out, requiring interpersonal comparisons to satisfy the preference principle in conjunction

⁴This is the approach moderately favored by Brun and Tungodden (2004) and formally explored by Sprumont (2011).

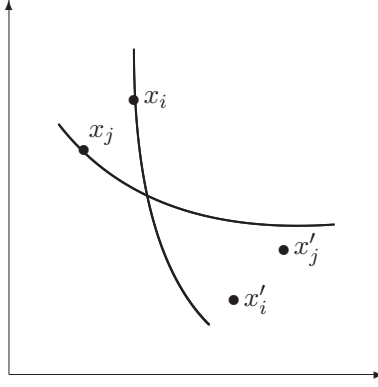


Figure 1: Incompatibility of the preference and the dominance principles

with the dominance principle restricted to a subset of X imposes a specific approach to interpersonal comparisons, namely, the *equivalence* approach. Let us explain this point. A set B is called a *monotone path* if $0 \in B$, B is unbounded and connected, and for all $x, x' \in B$, either $x \geq x'$ or $x \leq x'$. The equivalence approach consists in specifying a monotone path in X , and comparing (x, R) and (x', R') by the relative positions of the vectors x^* and x'^* from the path such that xIx^* and $x'I'x'^*$. This is illustrated in Figure 2, where (x, R) is declared inferior to (x', R') in this fashion, with the path given by the curve B .

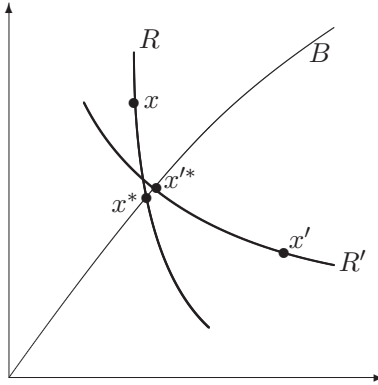


Figure 2: The equivalence approach

The following proposition, proven by Fleurbaey et al. (2009), states formally that if one wants to respect the preference principle and the restricted dominance principle, one necessarily has to follow the equivalence approach to rank situations (x, R) .

Proposition 1 (*Fleurbaey et al., 2009*). *Let B be a subset of X such that for every (x, R) there is x^* in B such that xIx^* . The dominance principle restricted to B , in conjunction with the preference principle, implies that $\succsim$ follows the equivalence approach with B as the monotone path.*

As explained in Fleurbaey (2009), the equivalence approach is not new. It basically boils down to the idea of money-metric utility that was popular in the 1980's (Deaton and Muellbauer, 1980; King, 1983). The theory of fair allocations has given it an original axiomatic justification, however. Moreover, it is clear that Proposition 1 does allow for different ways of making interpersonal comparisons, since it does not fix the monotone path (just as money-metric utilities depend on the choice of reference prices). The recent literature has shown that there may be good ethical reasons to choose a specific monotone path (Fleurbaey and Maniquet, 2011). Although some open questions remain, attractive solutions have been found for specific policy environments. One example will be described in the next subsection.

2.2 Social priorities

We now move beyond interpersonal comparisons and consider the issue of comparing social states.⁵ To fix ideas we work with a specific model in which the two relevant life dimensions are health and consumption (we follow Fleurbaey, 2005). An individual situation is $x_i = (c_i, h_i) \in \mathbb{R}_+ \times [0, 1]$. We keep the same assumptions about individual preferences R_i : they are assumed to be complete, monotonic and transitive. The fixed population is $N = \{1, \dots, n\}$ and an allocation is denoted $x_N = (x_1, \dots, x_n)$. The ranking

⁵ A closely related approach was already proposed in Willig (1981).

of allocations from the point of view of social welfare will be denoted $\bar{R}$ (with asymmetric and symmetric parts $\bar{P}$ and $\bar{I}$), and will be assumed to be transitive. Since we want this social ranking to depend on the profile of individual preferences $R_N = (R_1, \dots, R_n)$, it is really a function $\bar{R}(R_N)$, but the argument will sometimes be dropped to shorten notation.

It will turn out that one specific definition of equivalent income will play an essential role in what follows. This is found by choosing the monotone path B as the set of all points in X with $h = 1$.⁶ We then implicitly define the *healthy-equivalent income* $E(x_i, R_i)$ by

$$(E(x_i, R_i), 1) I_i(c_i, h_i), \quad (1)$$

i.e., as the level of consumption $E(x_i, R_i)$ that would make the individual indifferent between his current situation (c_i, h_i) and a situation in which she would be perfectly healthy and enjoy that level of consumption $E(x_i, R_i)$. The concept is illustrated in Figure 3.

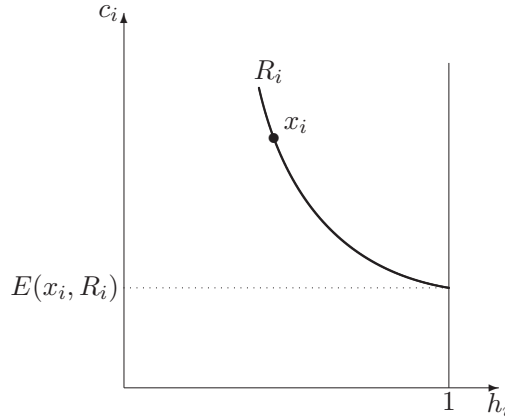


Figure 3: The healthy-equivalent income

Fleurbaey (2005) then formulates three requirements that can be imposed on the social ranking $\bar{R}(R_N)$. Universal quantifiers are omitted whenever the meaning of the axiom is clear.

⁶This set does not include 0, and one can add the bundles such that $c = 0$ to have a full path. But the result recalled below only deals with bundles that are at least as good as $(0, 1)$ for every individual.

Weak Pareto: If for all i , $x_i P_i x'_i$, then $x_N \bar{P}(R_N) x'_N$.

Independence: If for all i , and all $q \in X$, $x_i I_i q \Leftrightarrow x_i I'_i q$ and $x'_i I_i q \Leftrightarrow x'_i I'_i q$, then $x_N \bar{R}(R_N) x'_N$ if and only if $x_N \bar{R}(R'_N) x'_N$.

Pigou-Dalton: If there is i, j such that $h_i = h_j$ and $(c_i, h_i) = (c'_i - \delta, h'_i) > (c'_j + \delta, h'_j) = (c_j, h_j)$ for some $\delta > 0$ while $x'_k = x_k$ for all $k \neq i, j$, then $x_N \bar{R}(R_N) x'_N$ provided that either $R_i = R_j$ or $h_i = h_j = 1$.

The first axiom is standard. The second one defines the informational setting: it states that the social ranking of two allocations x and x' should be based only on information concerning the shape of the indifference curves through x_i and x'_i for all individuals i . This allows for richer information than Arrow's impossibility theorem, which would only consider individual pairwise preferences over x and x' . However, it also implies – in line with the preference principle – that information on subjective welfare levels (i.e. the cardinalization of the utility function) is irrelevant.

The third axiom introduces some egalitarianism in the space of resources (and obviously not in the space of subjective welfare levels). At first sight one could think that it would make sense to impose the restriction that a transfer of consumption from the rich to the poor increases social welfare, under the condition that the rich and the poor are at the same health level. However, Fleurbaey and Trannoy (2003) have shown that this requirement is incompatible with the Pareto condition, in a similar way as the dominance principle is incompatible with the preference principle. The third condition above therefore restricts the application of resource transfers to individuals with the same preferences or that have *perfect* health. This latter point is particularly important. The idea is that if two individuals are both perfectly healthy, then preferences “should not matter” in determining the desirability of an income transfer. With two individuals at the same mediocre health level, it may be legitimate for the rich to claim that he is in a worse situation because he cares more for his health. This reasoning is not at all convincing, however, if

he is in perfect health.

Fleurbaey (2005) then derived the following result:

Proposition 2 (*Fleurbaey, 2005*) *If the social ordering $\bar{R}(\cdot)$ satisfies weak Pareto, independence, and Pigou-Dalton, then $x_N \bar{P}(R_N) x'_N$ whenever $\min_i E(x_i, R_i) > \min_i E(x'_i, R_i)$.*

There are two noteworthy features about Proposition 2. First, taking perfect health as the reference in the Pigou-Dalton condition is one example of how ethical considerations can supplement the finding of Proposition 1. Indeed, imposing this condition “fixes” the choice of the monotone path that we described in the previous subsection – as illustrated in Figure 3, the healthy-equivalent income is one specific example of the more general idea of equivalent income. Second, although we did not impose an extreme form of egalitarianism, the combination of the Independence and Pigou-Dalton conditions imposes the maximin rule. This is reminiscent of a criticism that was raised against money-metric utilities by Blackorby and Donaldson (1988), namely, that in general they do not yield a quasi-concave social ordering over allocations. This problem disappears when the social ranking follows the maximin (or leximin) criterion. More on this can be found in Fleurbaey and Maniquet (2011).

3 Behavioral economics: shaking preferences?

The equivalent income approach, as described in the previous section, crucially depends on preferences. More specifically, the previous results are based on the assumption that there is a well-defined complete preference ordering. While this has been the traditional assumption in welfare economics, the belief in it has been shaken by recent findings from behavioral economics. It is not our point here to give a complete overview of all so-called behavioral anomalies that have been described in the literature, as there exist by now a lot of survey papers. Referring to just one of these that focuses on evidence from the field (Della Vigna, 2009), one can distinguish *non-standard preferences* (self-control

problems in an intertemporal setting; the influence of default options and the endowment effect), *non-standard beliefs* (economic agents overestimate their performance in tasks requiring ability, they expect small samples to exhibit large-sample statistical properties and they project their current preferences onto future periods) and *non-standard decision making* (the neglect or overweighting of information because of limited attention; the use of suboptimal heuristics for choices out of menu sets; excess impact of the beliefs of others; the possibly important role played by emotions such as mood and arousal). The findings from this literature suggest that preferences may not be regular – and that, even if such regular preferences did exist, choice behavior can in any case not be interpreted as the simple maximization of a fixed preference ordering. This raises difficult challenges for welfare economics.

One popular reaction in the behavioral literature has been to go back to experience utility (Frey and Stutzer, 2002; Kahneman et al., 1997; Kahneman and Sugden, 2005; Köszegi and Rabin, 2008; Layard, 2005). The intuition behind this is that, if people make mistakes, “decision utility” (the perceived utility on which decisions are based) and “experience utility” (the real after-decision utility) do no longer coincide and that in these circumstances it is better from the welfare point of view to focus on the “correct” outcomes. Yet, for the reasons explained earlier, this move back to welfarism is not an acceptable approach if one wants to respect preferences (see also Bernheim, 2009; Loewenstein and Ubel, 2008).

The alternative approach is to keep preferences as the ultimate criterion for evaluating social states, but to take into account that the preference relation that can be derived from choice behavior is not complete if choices are conflicting and context-dependent. Simply stated, one has to face the fact that in some circumstances x will be chosen from (x, y) and in other circumstances y . An interesting way to model this context-dependency has been proposed by Bernheim and Rangel (2009) and Salant and Rubinstein (2008). They introduce the concept of a generalized choice situation $G = (A, d)$ where A is the set of

elements from which a choice has to be made and d is an “ancillary condition” (in the terminology of Bernheim and Rangel, 2009) or a “frame” (in the terminology of Salant and Rubinstein, 2008). A standard choice situation would be fully characterized by A . Ancillary conditions (or frames) influence decisions but are (by definition) irrelevant for welfare. Examples of frames could be the specification of a default option or circumstances which lead to emotional arousal.⁷ The set of all generalized choice situations of interest is given by $\mathbb{C}$. The choice-correspondence for individual i is then given by $C_i(A, d) \subseteq A$ for all $(A, d) \in \mathbb{C}$. Its interpretation is obvious: $x \in C_i(G)$ is an object that individual i may choose when facing G . Note that it is very well possible that $C_i(A, d) \neq C_i(A, d')$. This is precisely how behavioral “anomalies” can be integrated in this framework.

If one has observations of individual behavior in different generalized choice situations, one can derive information about individual preferences over A . One possibility is to apply a structural model of behavior to explain the observations $C_i(A, d)$, i.e., to model how preferences together with frames determine choice. This structural model can then be used to derive a preference relation that is consistent with observed behavior conditional on the model used (Dalton and Ghosal, 2010; Rubinstein and Salant, 2009). The revealed preference relation is not necessarily complete. More importantly, it will depend on the specific behavioral model that is applied – and, very often it will be impossible to identify the correct model from the observations, in the sense that the outcomes of two different behavioral models (with different underlying preference relations) are observationally equivalent (in terms of choices).

Bernheim and Rangel (2009) therefore propose as an alternative what they call a “libertarian” approach, because it only uses information about choices. On this basis they define a series of incomplete welfare relations. The most attractive (and the one with

⁷Bernheim and Rangel (2009) and Salant and Rubinstein (2008) give many examples on how to cast the behavioral anomalies from the literature in the mould of generalized choice situations.

which they themselves work extensively) is

$$xP_i^*y \text{ iff, for all } (A, d) \in \mathbb{C} \text{ such that } x, y \in A, \text{ we have } y \notin C_i(A, d)$$

or: xP_i^*y iff, whenever both x and y are available, y is never chosen, and there is at least one $(A, d) \in \mathbb{C}$ such that $x, y \in A$ for which $x \in C_i(A, d)$.⁸ Using the relation P^* , x is a weak individual welfare optimum in A , if $\nexists y \in A$ such that yP^*x . This means that any object x chosen from a set A under some ancillary condition is a weak individual welfare optimum with the set A .

Bernheim and Rangel (2009) emphasize that their approach is only choice-based and does not assume the existence of an underlying preference relation. On the other hand, they are well aware that the relation P^* can be very coarse – and that in some generalized choice situations it is highly unlikely that a choice reveals something about welfare. The foreigner who is killed in a car accident in London because he did forget to look right when crossing the road does not reveal that he preferred being killed (unless there is other evidence in his life suggesting that he had suicidal tendencies). They therefore consider the possibility of refining P^* by using non-choice information to discard information from some generalized choice situations as “suspect”. In fact, consider two generalized choice domains $\mathbb{C}_1$ and $\mathbb{C}_2$ with $\mathbb{C}_1 \subset \mathbb{C}_2$, and two associated choice correspondences C_1 defined on $\mathbb{C}_1$ and C_2 defined on $\mathbb{C}_2$, with $C_1(G) = C_2(G)$ for all $G \in \mathbb{C}_1$. The welfare relation P_1^* obtained from $(\mathbb{C}_1, C_1)$ is then weakly finer than the welfare relation P_2^* obtained from $(\mathbb{C}_2, C_2)$, i.e. xP_2^*y implies xP_1^*y .

The decisions of what are “suspect” generalized choice situations will necessarily be to some extent subjective. Clear cases of biased information and misperception are perhaps easy to judge. Criteria of simplicity and coherence between behavior in different situations are already more debatable – and the use of self-reported preferences or in-

⁸The latter condition is always satisfied in Bernheim and Rangel (2009), because they assume that $((x, y), d) \in \mathbb{C}$ and the individual always selects some alternative in each generalized choice situation.

formed expert opinions (as advocated, e.g., by Beshears et al., 2008) is largely rejected by Bernheim (2009). In the extreme case, we may end up again in the Rubinstein-Salant (2009) approach where a specific model of behaviour is used to derive information about preferences from observed choice behavior – an approach that will lead to a finer relation, but, as mentioned before, requires one to choose between different models that may be observationally equivalent.

For our purposes, it is necessary to assume (contrary to Bernheim and Rangel, 2009) that there are preferences underlying choices – as these preferences will be the (normative) building block to arrive at an interpersonally comparable measure of well-being. Moreover, one may wonder how relevant behavioral anomalies are when one is concerned in the first place with “authentic” normative preferences, i.e. individual’s deep ideas about what is important in life. This is different from explaining and rationalizing behavior in real-life situations and it is likely that there is more scope for refinements in the former case. We will come back to this issue in the following section. Yet, as we will argue, there are very good reasons to think that even in this setting one will have to work with incomplete preference relations. The formal framework of Bernheim and Rangel (2009) and the characteristics of the relation P^* will therefore remain relevant in our setting too – even if we will interpret P^* as a preference relation.

Bernheim and Rangel (2009) have a counterexample showing that P^* is not necessarily transitive, but they show that P^* is acyclic. Imposing more structure on the space of alternatives may lead to P^* being transitive and, in fact, for almost all popular behavioral approaches, P^* is indeed transitive.

For applied welfare analysis, Bernheim and Rangel (2009) introduce natural counterparts of the concepts of compensating and equivalent variation. We will focus on the former. Let us assume that the generalized choice situation can be written as $((A(\alpha, m), d)$, where α is a vector of environmental parameters and m is a monetary transfer. Let us then consider a move from $((A(\alpha_0, 0), d_0)$ to $((A(\alpha_1, m), d_1)$. The compensating variation

is the smallest value of m such that for any $x \in C(A(\alpha_0, 0), d_0)$ and $y \in C(A(\alpha_1, m), d_1)$ the individual would be willing to choose y over x . In a setting with incomplete preferences, the latter sentence is ambiguous, however. We can consider the compensation to be sufficient when the new situation is unambiguously chosen over the old one or when the old situation is not unambiguously chosen over the new one. This leads to two notions of compensating variation. The first, CV^{high} , is the level of compensation m that solves

$$\inf \{m | y P^* x \text{ for all } m' \geq m, x \in C(A(\alpha_0, 0), d_0) \text{ and } y \in C(A(\alpha_1, m'), d_1)\}$$

The second, CV^{low} , is the level of compensation m that solves

$$\sup \{m | x P^* y \text{ for all } m' \leq m, x \in C(A(\alpha_0, 0), d_0) \text{ and } y \in C(A(\alpha_1, m'), d_1)\}$$

It is easy to see that $CV^{high} \geq CV^{low}$.

In a setting with many individuals, a move from x to y is a Pareto-improvement if $y P_i^* x$ for all i . If we do not define an interpersonally comparable concept of well-being, policy analysis remains restricted to looking for such Pareto-improvements. These will be very rare indeed and this usually motivates the use of the sum of compensating variations. While the compensating variation yields a specific measure of the welfare change for one individual, it is well known, however, that simply adding compensating variations is not an acceptable welfare criterion if one cares about the distribution and if one wants to avoid cyclic decisions (Blackorby and Donaldson, 1990). A setting with incomplete preferences does nothing to alleviate this criticism.

Yet, the intuition of looking for monetary equivalents is closely related to the intuition behind the equivalent income as introduced in the previous section. Moreover, the idea of upper and lower bounds is a natural way of coping with incomplete preferences. In the following sections we will show how these same intuitions can be useful to extend the (interpersonally comparable) equivalent income to a setting with incomplete preferences.

4 What about authentic preferences?

Are the findings of behavioral economics relevant within the normative perspective taken by the theory of fair social choice? An easy answer could be that they are not, since, as described before, the preferences we have in mind in this setting are people’s “authentic preferences” about what a good life is – not the preferences that are revealed in actual choice behavior. Authentic preferences are the preferences that would be revealed under “perfect” circumstances by well-informed individuals who consider all the relevant aspects of life and are not misled by cues of the environment. It could be argued that these preferences must necessarily be complete. If individuals are influenced by the framing of the choice problem or use imperfect information, this should be corrected. In the terminology introduced before: there is much scope and justification for refinements. Note that the argumentation about the refinements here can take an ethical twist: some ancillary conditions can be discarded even if they do not unambiguously involve biases, just because they are ethically unappealing. One example could be the importance of the reference situation – it is well known that people tend to focus on changes (gains and losses) rather than on the resulting final states and, moreover, that the feeling of loss looms larger than the feeling of gain. People will give larger subjective weight to avoiding the former than to experiencing the latter. One could argue that from the ethical point of view the status quo position should play a less prominent role, and definitely so if one is concerned about evaluating redistribution measures that are to the advantage of the poorest in society, i.e. where the rich lose and the poor gain. Of course, a justification of this kind of position ultimately rests on an ethical argumentation, that will not necessarily be accepted by everyone.⁹

This position that “authentic preferences” must necessarily be complete is too easy, however. First, even if one accepts that such authentic preferences do exist, there is still

⁹Noor (2010) suggests that choice-theoretic foundations for normative preferences can be built on the idea that the normative perspective is revealed when the agent is distanced from the consequences of his choice, an idea which is closely related to the veil of ignorance.

a challenge to discover them. Whether we simply ask people about their ideas, or we let them participate in choice experiments, or we start from real-life choices, or we ask them about their overall satisfaction with life – these “revealed” and “stated” preferences will be partly influenced by the way the choice situation is set up. Emotions and heuristics will also play a role in evaluating one’s overall life. It is possible that in a cold-blooded evaluation of the life cycle, emotions may receive less weight than in the heat of the moment. Yet, even if individuals are invited to abstract from immediate emotions, and even if they take a cognitive stance, they will still follow reason-based heuristics and the context will have an impact on their decisions (Shafir et al., 1993; Tversky and Simonson, 1993). Therefore, the idea of generalized choice situation is relevant also in this context – and the analyst will only be able to “derive” an incomplete (preference) relation P^* . While refinements are indeed possible, it is unlikely that there will be complete consensus about them or that we can refine until we ultimately end up with standard preferences. Moreover, the argument of observational equivalence between different behavioral models remains relevant in this setting also. Therefore, even if one accepts that individuals have well-defined authentic preferences, any cautious application will still have to admit our lack of knowledge about them. Working with incomplete preferences can then be seen as a kind of robustness check.

Second, one can even doubt that individuals have a complete preference relation over all possible lives. Indeed, this would imply that they can order states with which they are not at all familiar. The psychological uncertainty about preferences may be expected to grow when one goes further away from the actual situation. Yet, to calculate, e.g., healthy-equivalent incomes, we need information on the whole indifference curve. Is someone who has been chronically ill for a long time (or is handicapped since birth) able to evaluate trade-offs in a situation of (nearly) perfect health? Note that we are referring here to ordinal preferences and not to the effect of adaptation leading to smaller changes in subjective satisfaction levels.

The question of the existence of a “true” underlying well-defined and complete preference ordering is interesting from a philosophical point of view, but at this stage we can leave it open. For our purposes the first argument is already sufficient: in most cases, the analyst (or the policy maker) will have to work with incomplete preferences. The formal framework that was described in the previous section will therefore be relevant. However, the framework sketched in section 2 puts more structure on the decision problem than the abstract and general approach of Bernheim and Rangel (2009). Let us therefore describe the form taken by preference relations in our approach.

We assume that individual preferences take the form of *partial* orderings P^* defined on the set of relevant life dimensions X , with $X = \mathbb{R}_+^\ell$. The expression xP^*y means that x is strictly preferred to y . We assume that preferences are transitive (xP^*y and yP^*z implies xP^*z). As noted before, this is not a very strong requirement. Moreover, in our setting it also makes sense – certainly as a first approach – to assume that preferences are monotonic ($x > y$ implies xP^*y) and continuous: the sets $UC(x, P^*) = \{q \in X \mid qP^*x\}$ and $LC(x, P^*) = \{q \in X \mid xP^*q\}$ are open subsets of X and $x \in \partial UC(y, P^*)$ if and only if $y \in \partial LC(x, P^*)$.

Note that only assuming that $UC(x, P^*)$ and $LC(x, P^*)$ are open would not prevent some forms of discontinuity, namely, the appearance of “poles” at some points: one could have a sequence $x_n \rightarrow x$ such that $y \in LC(x_n, P^*)$ for all n but y belongs to the interior of $X \setminus LC(x, P^*)$. Under the additional condition that $x \in \partial UC(y, P^*)$ if and only if $y \in \partial LC(x, P^*)$, such phenomenon cannot occur because x is in the interior of $NC(y, P^*) = \{q \in X \mid \text{neither } qP^*y \text{ nor } yP^*q\}$ if and only if y is in the interior of $NC(x, P^*)$. Indeed, assume that x is in the interior of $NC(y, P^*)$. Then one cannot have $y \in \partial NC(x, P^*)$, which would require either $y \in \partial UC(x, P^*)$ or $y \in \partial LC(x, P^*)$. And one cannot have $y \notin NC(x, P^*)$, which would require either $y \in UC(x, P^*)$ or $y \in LC(x, P^*)$.

Under monotonicity and transitivity, $UC(x, P^*)$ is upper comprehensive, i.e., if $q \in UC(x, P^*)$, and $q' > q$, then $q' \in UC(x, P^*)$. Similarly, $LC(x, P^*)$ is lower comprehensive,

i.e., if $q \in LC(x, P^*)$ and $q' < q$, then $q' \in LC(x, P^*)$.

Under these conditions, it is therefore enough to know $NC(x, P^*)$ in order to know $UC(x, P^*)$ and $LC(x, P^*)$. One has

$$\begin{aligned} UC(x, P^*) &= \{q \in X \mid q \notin NC(x, P^*) \text{ and } \exists q' \in NC(x, P^*), q > q'\} \\ LC(x, P^*) &= \{q \in X \mid q \notin NC(x, P^*) \text{ and } \exists q' \in NC(x, P^*), q < q'\}. \end{aligned}$$

If xP^*y , then $LC(y, P^*) \subsetneq LC(x, P^*)$. First, one cannot have $LC(x, P^*) = LC(y, P^*)$ because $y \in LC(x, P^*)$ but $y \notin LC(y, P^*)$. Second, suppose that $LC(y, P^*) \subseteq LC(x, P^*)$ does not hold. Let $z \in LC(y, P^*) \setminus LC(x, P^*)$. One has yP^*z , which by transitivity implies xP^*z and therefore $z \in LC(x, P^*)$, a contradiction. Similarly, one shows that if xP^*y , then $UC(x, P^*) \subsetneq UC(y, P^*)$.

In the next sections it will be useful to address the following question. Consider two preferences $P^*, P^{*'}$ and two sets of points Q, Q' . Under what conditions does there exist a preference $P^{*''}$ such that for all $q \in Q$, $NC(q, P^{*''}) = NC(q, P^*)$ and for all $q' \in Q'$, $NC(q', P^{*''}) = NC(q', P^{*'})$? A sufficient condition is that for all $q \in Q$, $q' \in Q'$, $NC(q, P^*) \cap NC(q', P^{*'}) = \emptyset$. This condition is an extension of the notion of non-crossing indifference curves.

The resulting indifference curves in the two-dimensional case (with $x = (c, h)$) are represented in Figure 4. For convenience, we have drawn them in a strictly convex way, but convexity is not necessary for our analysis. Note that our assumptions imply that individuals have finer preferences when comparing close alternatives: this seems a very natural assumption.

5 Interpersonal comparisons with incomplete preferences

Let us now consider the issue of interpersonal comparisons, i.e., of ranking personal situations (x, P^*) in terms of well-being. As before, anonymity is assumed from the outset, so

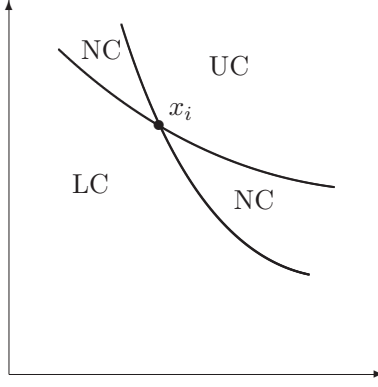


Figure 4: Indifference curves for incomplete preferences

that the identity of individuals is not part of the description of situations. The ranking is denoted $\succsim$ (with asymmetric and symmetric parts $\succ, \sim$) and is required to be transitive but not necessarily complete. The axioms and concepts from section 2 can now be adapted in a straightforward way (B is a subset of X):

Preference Principle: $(x, P^*) \succ (x', P^*)$ if $x P^* x'$.

Dominance Principle: $(x, P^*) \succsim (x', P^{*'})$ if $x \geq x'$; $(x, P^*) \succ (x', P^{*'})$ if $x \gg x'$.

Restricted Dominance Principle: For all $x, x' \in B$, $(x, P^*) \succsim (x', P^{*'})$ if $x \geq x'$;
 $(x, P^*) \succ (x', P^{*'})$ if $x \gg x'$.

The *equivalence approach* gathers the rankings $\succsim$ such that how to rank (x, P^*) and $(x', P^{*'})$ is fully determined by $UC(x, P^*) \cap B$, $LC(x, P^*) \cap B$ and $UC(x', P^{*'}) \cap B$, $LC(x', P^{*'}) \cap B$.

It is obvious that the preference principle and the dominance principle are incompatible, as the latter implies that $(x, P^*) \sim (x, P^{*'})$ for all x and all $P^*, P^{*'}$, making it impossible to take account of preferences. However, the (reformulated) preference and restricted dominance principles imply that the ordering $\succsim$ displays some aspects of the (reformulated) equivalence approach:

Proposition 3 *Let B be a subset of $\mathbb{R}_+^m$ such that for all (x, P^*) , $NC(x, P^*) \cap B \neq \emptyset$. If $\succsim$ satisfies the preference principle and the restricted dominance principle with respect to B , then B is a monotone path and $(x, P^*) \succ (x', P^{*'})$ whenever $LC(x, P^*) \cap UC(x', P^{*'}) \cap B \neq \emptyset$.*

This does not characterize the ranking fully. Assuming that B is a monotone path, every ranking such that:

- (i) $(x, P^*) \succ (x', P^{*'})$ whenever $LC(x, P^*) \cap UC(x', P^{*'}) \cap B \neq \emptyset$,
- (ii) $(x, P^*) \sim (x', P^{*'})$ whenever $x \in B$,
- (iii) $(x, P^*) \succ (x', P^*)$ whenever $x P^* x'$,

satisfies the preference principle and the restricted dominance principle. This allows many different possible rankings when none of these three situations applies. In particular, such rankings may involve the equivalence approach w.r.t. other paths, or non-equivalence approaches. As an example, compare the three cases in Figures 5-7. Proposition 3 allows us to say that $(x, P^*) \succ (x', P^{*'})$ in case A – but it does not allow a similar conclusion in cases B and C.

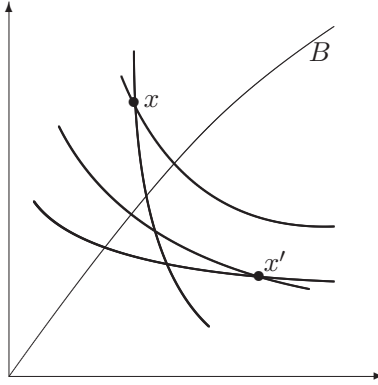


Figure 5: Case A

We get closer to the equivalence approach by adding additional requirements. One attractive condition is that one should avoid ranking an individual as better off than another individual when the available information about his situation is compatible with

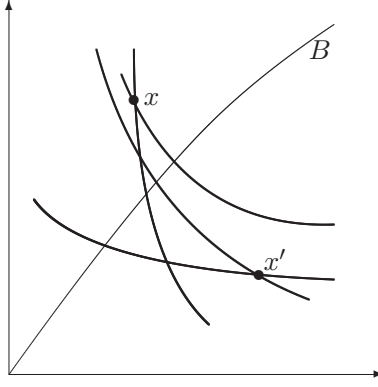


Figure 6: Case B

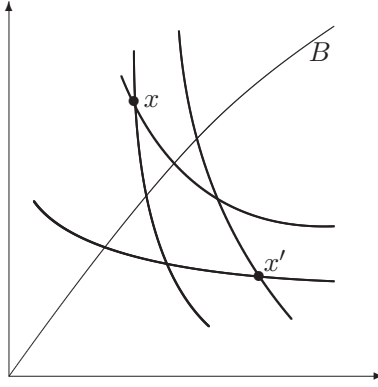


Figure 7: Case C

his being unambiguously worse off. This intuition is captured by the following Safety Principle:

Safety Principle: $(x, P^*) \succ (x', P^{*'})$ if there exists $\overline{P}^{*'} \supset P^{*'}$ such that for all $\overline{P}^* \supset P^*$,
 $(x, \overline{P}^*) \succ (x', \overline{P}^{*'})$.

Specifically, the safety principle says that if a refinement of one individual's preferences may reveal him to be worse off, then we should already consider him to be worse off. Of course, there are situations in which either individual can turn out to be worse off than the other when the information about both agents is refined. But the axiom deals with the case in which refining the information about one of them only may already determine

that he is worse off. The main motivation for this axiom is that, even though it does not preclude mistakes in interpersonal comparisons, it prevents the evaluator from missing a situation in which the worse-off is really badly off. If the evaluator is wrong about the worse-off in a pairwise comparison, the true worse-off is not as badly off as he could be if the mistake was in the opposite direction. Imposing it leads to the following proposition.

Proposition 4 *Let B be a subset of $\mathbb{R}_+^m$ such that for all (x, P^*) , $UC(x, P^*) \cap B \neq \emptyset$ and $LC(x, P^*) \cap B \neq \emptyset$. If $\succsim$ satisfies the preference principle and the restricted dominance principle with respect to B , then B is a monotone path and, under the safety principle, $(x, P^*) \succ (x', P^{*'})$ whenever $LC(x, P^*) \cap (\mathbb{R}_+^m \setminus LC(x', P^{*'})) \cap B \neq \emptyset$.*

Returning to the examples in the Figures, application of the safety principle now makes it possible to state that $(x, P^*) \succ (x', P^{*'})$ also in cases B and C. However, case C illustrates a limitation of the safety principle. Indeed, refining $P^{*'}$ to $\bar{P}^{*'}$ could also lead to the opposite configuration – implying that $(x, \bar{P}^*) \prec (x', \bar{P}^{*'})$. A variant of the safety principle is more cautious and checks that this cannot occur. Imposing it means that we can still draw conclusions in case B, but no longer in case C.

Super Safety Principle: $(x, P^*) \succ (x', P^{*'})$ if:

- (i) there exists $\bar{P}^{*'} \supset P^{*'}$ such that for all $\bar{P}^* \supset P^*$, $(x, \bar{P}^*) \succ (x', \bar{P}^{*'})$ or there exists $\bar{P}^* \supset P^*$ such that for all $\bar{P}^{*'} \supset P^{*'}$, $(x, \bar{P}^*) \succ (x', \bar{P}^{*'})$;
- (ii) there exists no $\bar{P}^{*'} \supset P^{*'}$ such that for all $\bar{P}^* \supset P^*$, $(x, \bar{P}^*) \prec (x', \bar{P}^{*'})$ and there exists no $\bar{P}^* \supset P^*$ such that for all $\bar{P}^{*'} \supset P^{*'}$, $(x, \bar{P}^*) \prec (x', \bar{P}^{*'})$.

Proposition 5 *Let B be a subset of $\mathbb{R}_+^m$ such that for all (x, P^*) , $UC(x, P^*) \cap B \neq \emptyset$ and $LC(x, P^*) \cap B \neq \emptyset$. If $\succsim$ satisfies the preference principle and the restricted dominance principle with respect to B , then B is a monotone path and, under the super safety principle, $(x, P^*) \succ (x', P^{*'})$ whenever $LC(x, P^*) \cap (\mathbb{R}_+^m \setminus LC(x', P^{*'})) \cap B \neq \emptyset$ and $UC(x', P^{*'}) \cap (\mathbb{R}_+^m \setminus UC(x, P^*)) \cap B \neq \emptyset$.*

It is clear that the propositions do not lead to a complete ranking of all possible personal situations (x, P^*) . Yet, in many cases they give an attractive operational criterion to define who is the worst off in a pairwise comparison. Moreover, our approach shares an attractive feature with the Bernheim-Rangel (2009) framework: if preferences are refined we come closer and closer to the standard equivalent income.

6 Incomplete preferences and social priorities

Let us now turn to the issue of comparing social states. As before, we will take the model of health and consumption used in Fleurbaey (2005), in which an individual situation is $x_i = (c_i, h_i) \in \mathbb{R}_+ \times [0, 1]$. However, we will now assume that (incomplete) individual preferences are denoted P_i^* and are assumed to be monotonic, transitive and to satisfy the continuity property introduced at the end of Section 4. For a given $x \in X$ and a given P^* , one can define the upper and the lower healthy-equivalent incomes

$$\begin{aligned} E^{\sup}(x, P^*) &= c^* \text{ such that } (c^*, 1) \in \partial UC(x, P^*), \\ E^{\inf}(x, P^*) &= c^* \text{ such that } (c^*, 1) \in \partial LC(x, P^*). \end{aligned}$$

We restrict attention to allocations x_N such that for all i , $E^{\inf}(x_i, P_i^*)$ and $E^{\sup}(x_i, P_i^*)$ are well defined.

Note that the previous propositions can be reformulated for the special case of healthy-equivalent incomes. Indeed, for this specific choice of monotone path, we can derive the simple operational criteria:

(preference principle, restricted dominance principle) $(x, P^*) \succ (x', P'^*)$ whenever $E^{\inf}(x, P^*) > E^{\sup}(x', P'^*)$

(preference principle, restricted dominance principle, safety principle) $(x, P^*) \succ (x', P'^*)$ whenever $E^{\inf}(x, P^*) > E^{\inf}(x', P'^*)$

(preference principle, restricted dominance principle, super safety principle) $(x, P^*) \succ (x', P'^*)$

$(x', P^{*'})$ whenever $E^{\inf}(x, P^*) > E^{\inf}(x', P^{*'})$ and $E^{\sup}(x, P^*) > E^{\sup}(x', P^{*'})$

Note the close relationship with the concepts of compensating variation CV^{high} and CV^{low} , as proposed by Bernheim and Rangel (2009). However, healthy-equivalent incomes yield an interpersonally comparable measure of well-being, i.e., an evaluation of the individual's overall personal situation, and not only a monetary evaluation of a change in this personal situation. They can therefore be used for social evaluation in cases where the distribution matters.

As before, we will now look for a ranking of allocations if we impose a Pigou-Dalton condition. We use the same notation $\bar{R}$ to denote this (transitive but not necessarily complete) ranking (with asymmetric and symmetric parts $\bar{P}$ and $\bar{I}$). This ranking is really a function $\bar{R}(P_N^*)$, but the argument will sometimes be dropped to shorten notations. Again, we can easily adapt the axioms that were introduced in section 2:

Weak Pareto: If for all i , $x_i P_i^* x'_i$, then $x_N \bar{P}(P_N^*) x'_N$.

Independence: If for all i , $NC(x_i, P_i^*) = NC(x_i, P_i^{*'})$ and $NC(x'_i, P_i^*) = NC(x'_i, P_i^{*'})$, then $x_N \bar{R}(P_N^*) x'_N$ if and only if $x_N \bar{R}(P_N^{*'}) x'_N$.

Pigou-Dalton: If there is i, j such that $h_i = h_j$ and $(c_i, h_i) = (c'_i - \delta, h'_i) > (c'_j + \delta, h'_j) = (c_j, h_j)$ for some $\delta > 0$ while $x'_k = x_k$ for all $k \neq i, j$, then $x_N \bar{R}(P_N^*) x'_N$ provided that either $P_i^* = P_j^*$ or $h_i = h_j = 1$.

We now have the following result.

Proposition 6 *If the social ordering $\bar{R}(\cdot)$ satisfies Weak Pareto, Independence, and Pigou-Dalton, then $x_N \bar{P}(P_N^*) x'_N$ whenever $\min_i E^{\inf}(x_i, P_i^*) > \min_i E^{\sup}(x'_i, P_i^*)$.*

It is possible to refine the ordering by adding a safety axiom once again.

Super Safety: If there is i and $P_i^{*'} \supset P_i^*$ such that $x_N \bar{P}(P_i^{*'}, P_{N \setminus \{i\}}^{*'}) x'_N$ for all $P_{N \setminus \{i\}}^{*'} \supset P_{N \setminus \{i\}}^*$, and for no j and $P_j^{*'} \supset P_j^*$ one has $x'_N \bar{P}(P_j^{*'}, P_{N \setminus \{j\}}^{*'}) x_N$ for all $P_{N \setminus \{j\}}^{*'} \supset P_{N \setminus \{j\}}^*$, then $x_N \bar{P}(P_N^*) x'_N$.

This axiom is similar to the Super Safety Principle of the previous section. It makes it possible to refine the ordering but not in a very simple way. Two observations.

First, the ordering defined by $x_N \bar{R}(P_N^*)x'_N$ iff $\min_i E^{\text{inf}}(x_i, P_i^*) \geq \min_i E^{\text{inf}}(x'_i, P_i^*)$ satisfies the four axioms.

Second, even if $\min_i E^{\text{inf}}(x_i, P_i^*) > \min_i E^{\text{inf}}(x'_i, P_i^*)$ and $\min_i E^{\text{sup}}(x_i, P_i^*) > \min_i E^{\text{sup}}(x'_i, P_i^*)$, the Super Safety axiom, in conjunction with the other three, does not imply that $x_N \bar{P}(P_N^*)x'_N$ in general. To see this, consider a case in which there is one agent i who is far worse-off than the others, so that the evaluation depends only on his preferences. If $x_i \in NC(x'_i, P_i^*)$, it may happen nonetheless that $E^{\text{inf}}(x_i, P_i^*) > E^{\text{inf}}(x'_i, P_i^*)$ and $E^{\text{sup}}(x_i, P_i^*) > E^{\text{sup}}(x'_i, P_i^*)$. This is compatible with finding $P_i^{*'} \supset P_i^*$ such that $x_i P_i^{*'} x'_i$ and $P_i^{*''} \supset P_i^*$ such that $x'_i P_i^{*''} x_i$. Therefore the Super Safety axiom has no bite in this case.

Of course, the social rankings that have been derived are incomplete. Yet, they are finer than the Pareto-ranking proposed by Bernheim and Rangel (2009) – and they allow to introduce distributional considerations in welfare analysis, even if one only uses information about ordinal and non-complete preferences. While it certainly would be worthwhile to explore further the potential contribution of imposing additional ethical requirements, the path that could be taken is clearly traced out.

7 Conclusion

We have argued in this paper that it is possible to define a concept of interpersonally comparable well-being that uses only information about ordinal preferences – even if these preferences are incomplete. Our paper therefore makes a contribution to two strands of the welfare economic literature. First, our introduction of incomplete preferences can be seen as an extension of the fair social choice approach. Second, we propose a method to define a normatively relevant concept of well-being as an extension of the Bernheim-Rangel (2009) approach to behavioral welfare economics. This makes it possible to go beyond

Pareto-efficiency and introduce distributional considerations into the welfare evaluation. Of course, for our approach to be meaningful it is necessary to assume that individuals do have preferences over different features of life. However, it is not necessary that these preferences are complete, nor that the analyst has perfect information about them.

The interpersonal comparisons and social rankings we derive are unavoidably incomplete. Yet, if one refines the individual preferences, one reaches the standard approach with equivalent incomes as a limiting case. Moreover a more complete social ranking can also be obtained by imposing additional normative requirements. Further work should look for a definition of acceptable and feasible refinements – or for the development of better methods to measure preferences.

The well-being concept we propose is very different from traditional “subjective utility” or “happiness”. We do not aim at measuring “true” happiness, but at formulating a concept that is meaningful for policy evaluation. Both the choice of the monotone path used in the equivalence approach and the choice of axioms to be imposed in the social evaluation exercise are essentially normative. This is not a weakness, but rather an advantage of the approach. When one aims at policy evaluation, it is better to make the underlying value judgments as open as possible. Having an informed debate about such value judgments in a formal model has always been the main objective of social choice theory.

References

- Bernheim, D. 2009. Behavioral welfare economics. *Journal of the European Economic Association* 7, no. 2-3: 267-319.
- Bernheim, D., and A. Rangel. 2009. Beyond revealed preference: choice theoretic foundations for behavioral welfare economics. *Quarterly Journal of Economics* 124, no. 1: 51-104.
- Beshears, J., J. Choi, D. Laibson, and B. Madrian. 2008. How are preferences

revealed? *Journal of Public Economics* 92, no. 8-9: 1787-94.

Blackorby, C., and D. Donaldson. 1988. Money metric utility: a harmless normalization? *Journal of Economic Theory* 46, no. 1: 120-129.

Blackorby, C., and D. Donaldson. 1990. A review article: the case against the use of the sum of compensating variations in cost-benefit analysis. *Canadian Journal of Economics* 23, no. 3: 471-94.

Brun, B., and B. Tungodden. 2004. Non-welfaristic theories of justice: Is the "intersection approach" a solution to the indexing impasse? *Social Choice and Welfare* 22, no. 1: 49-60.

d'Aspremont, C., and L. Gevers. 2002. Social welfare functionals and interpersonal comparability, in: K. Arrow, A. Sen, and K. Suzumura (eds.), *Handbook of Social Choice and Welfare*, Vol. 1, 459-541. Amsterdam: Elsevier.

Dalton, P. , and S. Ghosal. 2010. *Behavioral decisions and welfare*. Center Tilburg: Discussion Paper 2010-143.

Deaton, A., and J. Muellbauer. 1980. *Economics and consumer behaviour*. Cambridge: Cambridge University Press.

Della Vigna, S. 2009. Psychology and economics: evidence from the field. *Journal of Economic Literature* 47: 315-72.

Dworkin, R. 1981. What is equality? Part 1: Equality of welfare. *Philosophy and Public Affairs* 10: 185-246.

Fleurbaey, M. 2005. Health, wealth and fairness. *Journal of Public Economic Theory* 7, no. 2: 253-84.

Fleurbaey, M. 2009. Beyond GDP: the Quest for a Measure of Social Welfare. *Journal of Economic Literature* 47, no. 4: 1029-75.

Fleurbaey, M., and F. Maniquet. 2011. *A theory of fairness and social welfare*. Cambridge: Cambridge University Press.

Fleurbaey, M., E. Schokkaert, and K. Decancq. 2009. *What good is happiness?*

CORE : Discussion Paper 2009/17.

Fleurbaey, M., and A. Trannoy. 2003. The impossibility of a Paretian egalitarian. *Social Choice and Welfare* 21, no. 2: 243-63.

Frey, B., and A. Stutzer. 2002. What can economists learn from happiness research? *Journal of Economic Literature* 40, no. 2: 402-35.

Kahneman, D., and R. Sugden. 2005. Experienced utility as a standard of policy evaluation. *Environmental and Resource Economics* 32: 161-81.

Kahneman, D., P. Wakker, and R. Sarin. 1997. Back to Bentham? Explorations of experienced utility. *Quarterly Journal of Economics* 112: 375-406.

King, M. 1983. Welfare analysis of tax reforms using household data. *Journal of Public Economics* 23: 183-214.

Köszegi, B., and M. Rabin. 2008. Choices, situations, and happiness. *Journal of Public Economics* 92: 1821-32.

Layard, R. 2005. *Happiness: lessons from a new science*. Allan Lane.

Loewenstein, G., and P. Ubel. 2008. Hedonic adaptation and the role of decision and experienced utility in public policy. *Journal of Public Economics* 92: 1795-810.

Noor, J. 2010. *Subjective welfare*. Department of Economics, Boston University. Mimeo.

Rubinstein, A., and Y. Salant. 2009. *Eliciting welfare preferences from behavioral datasets*. Mimeo.

Salant, Y., and A. Rubinstein. 2008. (A,f): choice with frames. *Review of Economic Studies* 75: 1287-96.

Sen, A. 1970. *Collective choice and social welfare*. San Francisco: Holden Day.

———. 1985. *Commodities and capabilities*. Amsterdam: North-Holland.

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

Sprumont, Y. 2011. Resource egalitarianism with a dash of efficiency. *Journal of*

Economic Theory : forthcoming.

Tversky, A., and I. Simonson. 1993. Context-dependent preferences. *Management Science* 39, no. 10: 1179-89.

Willig, R. 1981. Social welfare dominance. *American Economic Review (Papers and Proceedings)* 71, no. 2: 200-204.

Appendix

Proof of Prop. 3 We first prove that B is a monotone path. As there is P^* such that $NC(0, P^*) = \{0\}$, and as B is such that $NC(0, P^*) \cap B \neq \emptyset$, necessarily $0 \in B$.

Let $z, z' \in B$ be such that neither $z \geq z'$ nor $z \leq z'$. There is P^* such that zP^*z' and $P^{*'} such that $z'P^{*'}z$. By the preference principle, $(z, P^*) \succ (z', P^*)$ and $(z', P^{*'}) \succ (z, P^{*'})$. By the restricted dominance principle, $(z, P^*) \sim (z, P^{*'})$ and $(z', P^*) \sim (z', P^{*'})$. This violates transitivity.$

The fact that for all (x, P^*) , $NC(x, P^*) \cap B \neq \emptyset$ then directly implies that B is unbounded and connected.

Let $(x, P^*), (x', P^{*'})$ be such that $LC(x, P^*) \cap UC(x', P^{*'}) \cap B \neq \emptyset$. Let $z \in LC(x, P^*) \cap UC(x', P^{*'}) \cap B$. By the preference principle, $(x, P^*) \succ (z, P^*)$ and $(z, P^{*'}) \succ (x', P^{*'})$. By the restricted dominance principle, $(z, P^*) \sim (z, P^{*'})$. By transitivity, $(x, P^*) \succ (x', P^{*'})$. ■

Proof of Prop. 4 Let $(x, P^*), (x', P^{*'})$ be such that $LC(x, P^*) \cap (\mathbb{R}_+^m \setminus LC(x', P^{*'})) \cap B \neq \emptyset$. As $LC(x, P^*)$ is open and lower comprehensive, while $\mathbb{R}_+^m \setminus LC(x', P^{*'})$ is closed and upper comprehensive, $LC(x, P^*) \cap (\mathbb{R}_+^m \setminus LC(x', P^{*'})) \cap B$ is not a singleton and there exist $z > z'$ in $LC(x, P^*) \cap (\mathbb{R}_+^m \setminus LC(x', P^{*'})) \cap B$. As z is not on the lower boundary of $\mathbb{R}_+^m \setminus LC(x', P^{*'})$, there is a refinement $\bar{P}^{*'} \supset P^{*'}$ such that $z \in UC(x', \bar{P}^{*'})$. For all refinements $\bar{P}^* \supset P^*$, $x\bar{P}^*z$ because xP^*z . By Prop. 3, $z \in LC(x, \bar{P}^*) \cap UC(x', \bar{P}^{*'}) \cap B$ implies that $(x, \bar{P}^*) \succ (x', \bar{P}^{*'})$. By the safety principle, $(x, P^*) \succ (x', P^{*'})$. ■

Proof of Prop. 5 Let $(x, P^*), (x', P^{*'})$ be such that $LC(x, P^*) \cap (\mathbb{R}_+^m \setminus LC(x', P^{*'})) \cap B \neq \emptyset$ and $UC(x', P^{*'}) \cap (\mathbb{R}_+^m \setminus UC(x, P^*)) \cap B \neq \emptyset$. Let $z \in LC(x, P^*) \cap (\mathbb{R}_+^m \setminus LC(x', P^{*'})) \cap B$ and $z' \in UC(x', P^{*'}) \cap (\mathbb{R}_+^m \setminus UC(x, P^*)) \cap B$ be such that z is not in the lower boundary of $\mathbb{R}_+^m \setminus LC(x', P^{*'})$ and z' is not in the upper boundary of $\mathbb{R}_+^m \setminus UC(x, P^*)$.

There is $\bar{P}^* \supset P^*$ such that for all $\bar{P}^{*'} \supset P^{*'}$, $z' \in LC(x, \bar{P}^*) \cap UC(x', \bar{P}^{*'}) \cap B$

(implying $(x, \bar{P}^*) \succ (x', \bar{P}^{*'})$ by Prop. 3). This implies that there is no $\bar{P}^{*'} \supset P^{*'}$ such that for all $\bar{P}^* \supset P^*$, $(x, \bar{P}^*) \prec (x', \bar{P}^{*'})$.

There is $\bar{P}^{*'} \supset P^{*'}$ such that for all $\bar{P}^* \supset P^*$, $z \in LC(x, \bar{P}^*) \cap UC(x', \bar{P}^{*'}) \cap B$ (implying $(x, \bar{P}^*) \succ (x', \bar{P}^{*'})$ by Prop. 3). This implies that there is no $\bar{P}^* \supset P^*$ such that for all $\bar{P}^{*'} \supset P^{*'}$, $(x, \bar{P}^*) \prec (x', \bar{P}^{*'})$.

By the super safety principle, $(x, P^*) \succ (x', P^{*'})$. ■

Proof of Prop. 6 : Let x_N, x'_N be such that $\min_i E^{\inf}(x_i, P_i^*) > \min_i E^{\sup}(x'_i, P_i^*)$.

Figure 8 illustrates the proof.

There exist $\hat{x}_N, \hat{x}'_N$ such that for all $i \in N$, $\hat{h}_i = \hat{h}'_i = 1$, $x_i P_i^* \hat{x}_i$, $\hat{x}'_i P_i^* x'_i$, and

$$\min_i E^{\inf}(x_i, P_i^*) > \min_i \hat{c}_i > \min_i \hat{c}'_i > \min_i E^{\sup}(x'_i, P_i^*).$$

Moreover, one can construct $\hat{x}_N, \hat{x}'_N$ so that there is a unique i_0 such that $\hat{c}_{i_0} = \min_i \hat{c}_i$ and $\hat{c}'_{i_0} = \min_i \hat{c}'_i$, and so that $\hat{x}'_i P_i^* \hat{x}_i$ for all $i \neq i_0$.

There exist $\bar{x}_{i_0}, \bar{x}'_{i_0}$ such that $\bar{h}_{i_0} = \bar{h}'_{i_0} < 1$ and $\hat{x}_{i_0} P_{i_0}^* \bar{x}_{i_0} P_{i_0}^* \bar{x}'_{i_0} P_{i_0}^* \hat{x}'_{i_0}$. For each $i \neq i_0$, let $\bar{x}_i, \bar{x}'_i$ be such that $\bar{h}_i = \bar{h}'_i = \bar{h}_{i_0}$, $\bar{x}'_i P_i^* \hat{x}'_i$ and $\bar{c}'_i - \bar{c}_i = (\bar{c}_{i_0} - \bar{c}'_{i_0}) / (n - 1)$.

There exist $\bar{x}'''_{i_0}$ such that $\bar{h}'''_{i_0} = 1$ and $\hat{x}_{i_0} P_{i_0}^* \bar{x}'''_{i_0} P_{i_0}^* \bar{x}_{i_0}$. For each $i \neq i_0$, let $\bar{x}''_i, \bar{x}'''_i$ be such that $\bar{h}''_i = \bar{h}'''_i = 1$, $\hat{x}_i > \bar{x}''_i > \bar{x}'''_i > \hat{x}_{i_0}$ and $\bar{c}'''_i - \bar{c}''_i = (\hat{c}_{i_0} - \bar{c}'''_{i_0}) / (2(n - 1))$. Let $\bar{x}''_{i_0} = (\hat{x}_{i_0} + \bar{x}'''_{i_0}) / 2$. One has $\bar{c}'''_i - \bar{c}''_i = (\bar{c}''_{i_0} - \bar{c}'''_{i_0}) / (n - 1)$.

Let $P_{i_0}^{*'} = P_{i_0}^*$ and for $i \neq i_0$, let $P_i^{*'}$ be such that $\bar{x}'_i P_i^{*'} \hat{x}'_i$, $\bar{x}'''_i P_i^{*'} \bar{x}_i$, $NC(\bar{x}'_i, P_i^{*'}) \cap NC(\hat{x}'_i, P_i^*) = \emptyset$, $NC(\hat{x}_i, P_i^*) \cap NC(\bar{x}''_i, P_i^{*'}) = \emptyset$, $NC(\bar{x}_i, P_i^{*'}) \cap NC(\bar{x}_{i_0}, P_{i_0}^{*'}) = \emptyset$.

For $i \neq i_0$, let $P_i^{*''}$ be such that $NC(\bar{x}'_i, P_i^{*''}) = NC(\bar{x}'_i, P_i^{*'})$, $NC(\bar{x}_i, P_i^{*''}) = NC(\bar{x}_i, P_i^{*'})$, and for all x such that $\bar{x}'_{i_0} \leq x \leq \bar{x}_{i_0}$, $NC(x, P_i^{*''}) = NC(x, P_{i_0}^{*'})$.

Number the agents $i \neq i_0$ from 1 to $n - 1$. By Pigou-Dalton,

$$(\bar{x}_1, \bar{x}'_2, \dots, \bar{x}'_{n-1}, \bar{x}'_{i_0} + \bar{x}'_1 - \bar{x}_1) \bar{R}(P_1^{*''}, P_2^{*'}, \dots, P_{n-1}^{*'}, P_1^{*''}) \bar{x}'_N,$$

and by independence,

$$(\bar{x}_1, \bar{x}'_2, \dots, \bar{x}'_{n-1}, \bar{x}'_{i_0} + \bar{x}'_1 - \bar{x}_1) \bar{R}(P_N^*) \bar{x}'_N.$$

Repeating this argument for agent 2, one obtains

$$(\bar{x}_1, \bar{x}_2, \bar{x}'_3, \dots, \bar{x}'_{n-1}, \bar{x}'_{i_0} + 2(\bar{x}'_1 - \bar{x}_1)) \bar{R}(P_N^*) \bar{x}'_N.$$

After applying this argument also to $i = 3, \dots, n-1$, and noting that $(n-1)(\bar{x}'_1 - \bar{x}_1) = \bar{x}_{i_0} - \bar{x}'_{i_0}$, one obtains $\bar{x}_N \bar{R}(P_N^*) \bar{x}'_N$.

By weak Pareto, $\bar{x}_N \bar{P}(P_N^*) \bar{x}_N$. By $n-1$ applications of Pigou-Dalton, $\bar{x}_N \bar{R}(P_N^*) \bar{x}'''_N$.

By transitivity, $\bar{x}_N \bar{P}(P_N^*) \bar{x}'_N$.

Let $P_{i_0}^{*'''} = P_{i_0}^*$ and for $i \neq i_0$, let $P_i^{*'''}$ be such that $NC(\bar{x}'_i, P_i^{*'''}) = NC(\bar{x}'_i, P_i^*)$, $NC(\bar{x}''_i, P_i^{*'''}) = NC(\bar{x}''_i, P_i^*)$, and $NC(\hat{x}_i, P_i^{*'''}) = NC(\hat{x}_i, P_i^*)$, $NC(\hat{x}'_i, P_i^{*'''}) = NC(\hat{x}'_i, P_i^*)$.

By independence, $\bar{x}_N \bar{P}(P_N^{*'''}) \bar{x}'_N$.

By weak Pareto, $\hat{x}_N \bar{P}(P_N^{*'''}) \bar{x}''_N$ and $\bar{x}'_N \bar{P}(P_N^{*'''}) \hat{x}'_N$. By transitivity, $\hat{x}_N \bar{P}(P_N^{*'''}) \hat{x}'_N$.

By independence, $\hat{x}_N \bar{P}(P_N^*) \hat{x}'_N$. By weak Pareto, $x_N \bar{P}(P_N^*) \hat{x}_N$ and $\hat{x}'_N \bar{P}(P_N^*) x'_N$. By transitivity, $x_N \bar{P}(P_N^*) x'_N$.

■

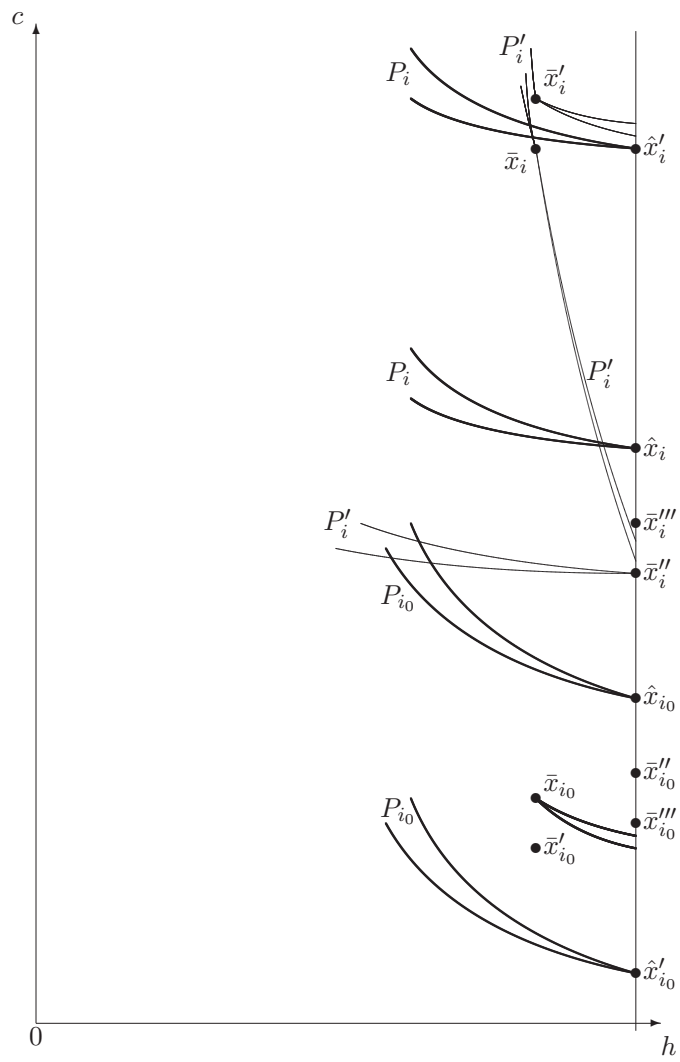


Figure 8: Illustration of the proof of Prop. 6.

Recent titles

CORE Discussion Papers

- 2011/1. Yu. NESTEROV. Random gradient-free minimization of convex functions.
- 2011/2. Olivier DEVOLDER, François GLINEUR and Yu. NESTEROV. First-order methods of smooth convex optimization with inexact oracle.
- 2011/3. Luc BAUWENS, Gary KOOP, Dimitris KOROBILIS and Jeroen V.K. ROMBOUTS. A comparison of forecasting procedures for macroeconomic series: the contribution of structural break models.
- 2011/4. Taoufik BOUEZMARNI and Sébastien VAN BELLEGEM. Nonparametric Beta kernel estimator for long memory time series.
- 2011/5. Filippo L. CALCIANO. The complementarity foundations of industrial organization.
- 2011/6. Vincent BODART, Bertrand CANDELON and Jean-François CARPANTIER. Real exchanges rates in commodity producing countries: a reappraisal.
- 2011/7. Georg KIRCHSTEIGER, Marco MANTOVANI, Ana MAULEON and Vincent VANNETELBOSCH. Myopic or farsighted? An experiment on network formation.
- 2011/8. Florian MAYNERIS and Sandra PONCET. Export performance of Chinese domestic firms: the role of foreign export spillovers.
- 2011/9. Hiroshi UNO. Nested potentials and robust equilibria.
- 2011/10. Evgeny ZHELOBODKO, Sergey KOKOVIN, Mathieu PARENTI and Jacques-François THISSE. Monopolistic competition in general equilibrium: beyond the CES.
- 2011/11. Luc BAUWENS, Christian HAFNER and Diane PIERRET. Multivariate volatility modeling of electricity futures.
- 2011/12. Jacques-François THISSE. Geographical economics: a historical perspective.
- 2011/13. Luc BAUWENS, Arnaud DUFAYS and Jeroen V.K. ROMBOUTS. Marginal likelihood for Markov-switching and change-point GARCH models.
- 2011/14. Gilles GRANDJEAN. Risk-sharing networks and farsighted stability.
- 2011/15. Pedro CANTOS-SANCHEZ, Rafael MONER-COLONQUES, José J. SEMPERE-MONERRIS and Oscar ALVAREZ-SANJAIME. Vertical integration and exclusivities in maritime freight transport.
- 2011/16. Géraldine STRACK, Bernard FORTZ, Fouad RIANE and Mathieu VAN VYVE. Comparison of heuristic procedures for an integrated model for production and distribution planning in an environment of shared resources.
- 2011/17. Juan A. MAÑEZ, Rafael MONER-COLONQUES, José J. SEMPERE-MONERRIS and Amparo URBANO. Price differentials among brands in retail distribution: product quality and service quality.
- 2011/18. Pierre M. PICARD and Bruno VAN POTTELSBERGHE DE LA POTTERIE. Patent office governance and patent system quality.
- 2011/19. Emmanuelle AURIOL and Pierre M. PICARD. A theory of BOT concession contracts.
- 2011/20. Fred SCHROYEN. Attitudes towards income risk in the presence of quantity constraints.
- 2011/21. Dimitris KOROBILIS. Hierarchical shrinkage priors for dynamic regressions with many predictors.
- 2011/22. Dimitris KOROBILIS. VAR forecasting using Bayesian variable selection.
- 2011/23. Marc FLEURBAEY and Stéphane ZUBER. Inequality aversion and separability in social risk evaluation.
- 2011/24. Helmuth CREMER and Pierre PESTIEAU. Social long term care insurance and redistribution.
- 2011/25. Natali HRITONENKO and Yuri YATSENKO. Sustainable growth and modernization under environmental hazard and adaptation.
- 2011/26. Marc FLEURBAEY and Erik SCHOKKAERT. Equity in health and health care.
- 2011/27. David DE LA CROIX and Axel GOSSERIES. The natalist bias of pollution control.
- 2011/28. Olivier DURAND-LASSERVE, Axel PIERRU and Yves SMEERS. Effects of the uncertainty about global economic recovery on energy transition and CO₂ price.
- 2011/29. Ana MAULEON, Elena MOLIS, Vincent J. VANNETELBOSCH and Wouter VERGOTE. Absolutely stable roommate problems.

Recent titles

CORE Discussion Papers - continued

- 2011/30. Nicolas GILLIS and François GLINEUR. Accelerated multiplicative updates and hierarchical algorithms for nonnegative matrix factorization.
- 2011/31. Nguyen Thang DAO and Julio DAVILA. Implementing steady state efficiency in overlapping generations economies with environmental externalities.
- 2011/32. Paul BELLEFLAMME, Thomas LAMBERT and Armin SCHWIENBACHER. Crowdfunding: tapping the right crowd.
- 2011/33. Pierre PESTIEAU and Gregory PONTIERE. Optimal fertility along the lifecycle.
- 2011/34. Joachim GAHUNGU and Yves SMEERS. Optimal time to invest when the price processes are geometric Brownian motions. A tentative based on smooth fit.
- 2011/35. Joachim GAHUNGU and Yves SMEERS. Sufficient and necessary conditions for perpetual multi-assets exchange options.
- 2011/36. Miguel A.G. BELMONTE, Gary KOOP and Dimitris KOROBILIS. Hierarchical shrinkage in time-varying parameter models.
- 2011/37. Quentin BOTTON, Bernard FORTZ, Luis GOUVEIA and Michael POSS. Benders decomposition for the hop-constrained survivable network design problem.
- 2011/38. J. Peter NEARY and Joe THARAKAN. International trade with endogenous mode of competition in general equilibrium.
- 2011/39. Jean-François CAULIER, Ana MAULEON, Jose J. SEMPERE-MONERRIS and Vincent VANNETELBOSCH. Stable and efficient coalitional networks.
- 2011/40. Pierre M. PICARD and Tim WORRALL. Sustainable migration policies.
- 2011/41. Sébastien VAN BELLEGEM. Locally stationary volatility modelling.
- 2011/42. Dimitri PAOLINI, Pasquale PISTONE, Giuseppe PULINA and Martin ZAGLER. Tax treaties and the allocation of taxing rights with developing countries.
- 2011/43. Marc FLEURBAEY and Erik SCHOKKAERT. Behavioral fair social choice.

Books

- J. HINDRIKS (ed.) (2008), *Au-delà de Copernic: de la confusion au consensus ?* Brussels, Academic and Scientific Publishers.
- J-M. HURIOT and J-F. THISSE (eds) (2009), *Economics of cities*. Cambridge, Cambridge University Press.
- P. BELLEFLAMME and M. PEITZ (eds) (2010), *Industrial organization: markets and strategies*. Cambridge University Press.
- M. JUNGER, Th. LIEBLING, D. NADDEF, G. NEMHAUSER, W. PULLEYBLANK, G. REINELT, G. RINALDI and L. WOLSEY (eds) (2010), *50 years of integer programming, 1958-2008: from the early years to the state-of-the-art*. Berlin Springer.
- G. DURANTON, Ph. MARTIN, Th. MAYER and F. MAYNERIS (eds) (2010), *The economics of clusters – Lessons from the French experience*. Oxford University Press.
- J. HINDRIKS and I. VAN DE CLOOT (eds) (2011), *Notre pension en héritage*. Itinera Institute.
- M. FLEURBAEY and F. MANIQUET (eds) (2011), *A theory of fairness and social welfare*. Cambridge University Press.
- V. GINSBURGH and S. WEBER (eds) (2011), *How many languages make sense? The economics of linguistic diversity*. Princeton University Press.

CORE Lecture Series

- D. BIENSTOCK (2001), Potential function methods for approximately solving linear programming problems: theory and practice.
- R. AMIR (2002), Supermodularity and complementarity in economics.
- R. WEISMANTEL (2006), Lectures on mixed nonlinear programming.