External Norms and Rationality of Choice

Walter Bossert

University of Montréal

Kotaro Suzumura

Waseda University

Acknowledgments. Thanks are due to Nick Baigent, Wulf Gaertner, Prasanta Pattanaik, Amartya Sen and Yongsheng Xu for discussions over the years on this and related issues. We are grateful to Bhaskar Dutta, Donald Saari, Thomas Schwartz as well as an editor and two referees of this Journal for their helpful comments. Financial support through grants from the Fonds de Recherche sur la Société et la Culture of Québec, a Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology of Japan and the Social Sciences and Humanities Research Council of Canada is gratefully acknowledged.

Abstract. Ever since Sen criticized the notion of internal consistency of choice, there exists a widespread perception that the standard rationalizability approach to the theory of choice has difficulties in coping with the existence of external norms. We introduce a concept of *norm-conditional rationalizability* and show that external norms can be made compatible with the methods underlying the traditional rationalizability approach. To do so, we characterize norm-conditional rationalizability by means of suitable modifications of revealed preference axioms that are well established in the theory of rational choice on general domains as analyzed in contributions by Richter and Hansson, for example. We compare our approach to alternative suggestions that have appeared in response to Sen's criticisms, and we discuss its links to Sen's notion of *self-imposed choice constraints.*

1 Introduction

In his Presidential Address to the Econometric Society, Sen (1993) constructed an argument which seems to go against *a priori* imposition of the requirement of internal consistency of choice such as the *weak axiom of revealed preference* (Samuelson 1938; 1947; 1948; 1950), the *strong axiom of revealed preference* (Houthakker 1950), *Arrow's axiom of choice consistency* (Arrow 1959), and Sen's *condition* α (Sen 1971), and investigated the logical implications of eschewing these internal choice consistency requirements. On the face of it, Sen's argument may seem to go squarely against the theory of rationalizability due to Arrow (1959), Richter (1966; 1971), Hansson (1968), Sen (1971), Suzumura (1976a; 1977; 1983), Bossert, Sprumont and Suzumura (2005; 2006) and many others. This paper argues that there is in fact a way of constructing a bridge between Sen's criticism against the requirement of internal consistency of choice and the traditional framework of rational choice. For the sake of pursuing this salvage activity, we introduce a novel concept of *norm-conditional rationalizability,* which helps us establish the peaceful co-existence of the rationalizability approach and Sen's criticism against the requirement of internal consistency of choice.

The gist of our approach is simple and intuitive. In the presence of external norms, some alternatives, which are physically feasible, may not necessarily be choosable without violating the given norm. It should be clear that this additional constraint imposed by the prevailing norm should be taken into due consideration in defining the rationality, or the lack thereof, of observed choice behavior. The concept of norm-conditional rationalizability is nothing other than a modification of the traditional concept of rationalizability that takes into account additional restrictions imposed by the norm. Along with this modification, the standard revealed preference relations must be amended in an analogous manner. Once we implement these modifications, we establish the theory of norm-conditional rationalizability, which is a natural generalization of the traditional theory of rationalizability. Indeed, the novel theory of norm-conditional rationalizability boils down to the traditional theory of rationalizability in the absence of external norms.

More precisely, a norm can be expressed by specifying pairs of alternatives and feasible sets containing the requisite alternative such that the choice of the alternative from this set is prohibited by the norm under considerations. For instance, consider the following example due to Sen (1993: 501). "Suppose [a] person faces a choice at a dinner table between having the last remaining apple in the fruit basket (y) and having nothing instead (x), forgoing the nice-looking apple. She decides to behave decently and picks nothing (x), rather than the one apple (y). If, instead, the basket had contained two apples, and she had encountered the choice between having nothing (x), having one nice apple (y) and having another nice one (z), she could reasonably enough choose one (y), without violating any rule of good behavior. The presence of another apple (z) makes one of the two apples decently choosable, but this combination of choices would violate the standard consistency conditions ... even though there is nothing particularly 'inconsistent' in this pair of choices ... ." The external norm invoked by Sen can be described as the maxim 'do not choose the last available apple' and a norm expressing such a rule of conduct can be defined by excluding the choice of a single apple from a feasible set that consists of one apple only and, at the same time, permitting the choice of one apple if there are at least two available. We will discuss this and other examples in more detail once we have introduced the requisite formal definitions.

Apart from this introduction, this paper consists of six sections. In Section 2, we examine Sen's criticism directed towards conditions expressing internal choice consistency properties and motivate our concept of external norms and that of norm-conditional rationalizability. Section 3 introduces the basic idea underlying our way of modelling norms. In Section 4, we state our main results regarding norm-conditional rationalizability. Section 5 discusses alternative formulations of the concept of external norms and compares them with the setup utilized in this paper. We will also make a few brief remarks on some earlier attempts to accommodate Sen's (1993) criticism on the internal consistency of choice. Section 6 concludes and Section 7 gathers the proofs of our results.

2 Sen's Criticism of Internal Consistency of Choice

Let C be a *choice function* that specifies, for any admissible non-empty set S of feasible alternatives, a non-empty subset $C(S)$ of S. In the standard case, there are no further restrictions on the choice set $C(S)$; in the norm-respecting cases to be considered in this paper, the choice set has to be consistent with the external norms imposed. A more precise definition is provided shortly. Note that the standard case is the special case of the norm-respecting framework in which there are no norms. For that reason, we simply refer to C as a choice function without specifying the assumption that it is norm-respecting: there is no need to make a distinction because the traditional setting is a special case.

Sen (1993: 500) poses the following question: "[C]an a set of choices really be seen as consistent *or* inconsistent on purely internal grounds *without* bringing in something *external* to choice, such as the underlying objectives or values that are pursued or acknowledged by choice?" To bring his point into clear relief, Sen invites us to examine the following two choices:

$C(\{x, y\}) = \{x\}$ and $C(\{x, y, z\}) = \{y\}$.

As Sen rightly points out, this pair of choices violates most of the standard choice consistency conditions including the weak and the strong axioms of revealed preference, Arrow's axiom of choice consistency, and Sen's condition α. It is arguable and indeed Sen (1993: 501) argues that this seeming inconsistency can be easily resolved if only we know more about the person's choice situation: as mentioned in the Introduction, the person making these choices may simply respect the external norm 'do not choose the last available apple.'

On the face of it, Sen's argument to this effect may seem to go squarely against the theory of rationalizability *à la* Arrow (1959), Richter (1966; 1971), Hansson (1968), Sen (1971), Suzumura (1976a) and many others, where the weak axiom of revealed preference is a necessary condition for rationalizability.

The standard theory of revealed preference has its origins in the analysis of consumer behavior where feasible sets are given by budget sets, pioneered by authors such as Samuelson (1938; 1947; 1948, 1950), Houthakker (1950) and many others. However, much more general approaches have been developed in the meantime, particularly those due to Richter (1966; 1971), Hansson (1968), Suzumura (1976a; 1977; 1983) and others who explored the traditional notion of rationalizability without any domain restrictions whatsoever. It is this general theory of rationalizability that we modify so as to develop a new concept of *norm-conditional rationalizability* and build a bridge between rationalizability and Sen's criticism. In essence, what emerges from this modification is the peaceful co-existence of a norm-conditional notion of rationalizability and Sen's elaborate criticism of internal consistency conditions. Although Sen's suggestion to the effect that the rationality of choice behavior, or the lack thereof, cannot be judged only by means of the internal structure of choices made is well taken, it turns out that we can modify the axioms of revealed preference theory in such a way as to provide an axiomatization of choices under external norms in terms of suitably modified revealed preference axioms.

As an auxiliary step, we introduce a model of choice where external norms are taken into consideration by specifying all pairs consisting of a feasible set and an element of this set with the interpretation that this element is prohibited from being chosen from this set by the relevant system of external norms. Norm-conditional rationalizability then requires the existence of a preference relation such that, for each feasible set in the domain of the choice function, the chosen elements are at least as good as all elements in the set *except* for those that are prohibited by the external norm. This approach is very general because no restrictions are imposed on how the system of external norms comes about—any specification of a set of pairs as described above is possible. Needless to say, we do not by any means suggest that any arbitrary system of norms thus specified is desirable; clearly, what we advocate is a *method* to incorporate any norm into a model of choice without completely eschewing all notions of traditional rationality altogether. In fact, the traditional model of rational choice is included as a special case—the case that obtains if the set of prohibited pairs is empty.

We note in passing that the 'do not choose the last available apple' example is not the only one Sen (1993) uses to criticize internal choice consistency properties. A second example he uses to call into question the imposition of internal choice consistency conditions focuses on the *epistemic value* of a feasible set. We do not pursue this issue further in this paper and refer the reader to Bossert (2001) for a detailed discussion and an attempt to link this phenomenon to standard rationalizability properties within a model of decision making under complete uncertainty.

3 Modelling External Norms

A choice situation is described by a feasible set S of alternatives, where S is a non-empty subset of the non-empty universal set X. External norms such as those discussed in the Introduction can be expressed by identifying feasible sets and alternatives that are not to be chosen from these feasible sets. For example, suppose there is a feasible set $S = \{x, y\}$, where x stands for selecting nothing and y stands for selecting a single apple. Now consider the feasible set $T = \{x, y, z\}$ where there are two (identical) apples y and z available. The external norm not to take the last apple can easily and intuitively be expressed by requiring that the choice of y from S is excluded, whereas the choice of y (or z) from T is perfectly acceptable. In general, norms of that nature can be expressed by identifying all pairs (S, w), where $w \in S$, such that w is not supposed to be chosen from the feasible set S. To that end, we use a set $\mathcal{N}$, to be interpreted as the set of all pairs (S, w) of a feasible set S and an element w of S such that the choice of w from S is prevented by the external norm under consideration.

More formally, suppose $\mathcal{X}$ is the power set of X excluding the empty set. A choice function with a non-empty domain $\Sigma \subseteq \mathcal{X}$ is a mapping $C\colon \Sigma \to \mathcal{X}$ such that, for all $S \in \Sigma$, $C(S) \subseteq S \setminus \{z \in S \mid (S, z) \in \mathcal{N}\}$. Let $C(\Sigma)$ denote the image of Σ under C, that is, $C(\Sigma) = \cup_{S \in \Sigma} C(S)$. To ensure that the non-emptiness of the chice set does not conflict with the restrictions imposed by the norm $\mathcal{N}$, we require $\mathcal{N}$ to be such that, for all $S \in \Sigma$, there exists $x \in S$ satisfying $(S, x) \notin \mathcal{N}$. The set of all possible norms satisfying this restriction is denoted by $\mathbf{N}$. If $\mathcal{N}$ is empty (that is, there are no external norms), we obtain the standard model of choice.

Returning to Sen's example involving the norm 'do not choose the last available apple,' we can, for instance, define the universal set $X = \{x, y, z\}$, the domain $\Sigma = \{S, T\} \subsetneq \mathcal{X}$ with $S = \{x, y\}$ and $T = \{x, y, z\}$, and the external norm described by the set $\mathcal{N} = \{(S, y)\}$. Thus, the external norm requires that $y \notin C(S)$ but no restrictions are imposed on the choice $C(T)$ from the set T—that is, this external norm represents the requirement that the last available apple should not be chosen.

The notion of rationality explored in this paper is conditional on a system of external norms $\mathcal{N} \in \mathbf{N}$ as introduced above. In contrast with the classical model of rational choice, an element x that is chosen by a choice function C from a feasible set $S \in \Sigma$ need not be considered at least as good as *all* elements of S by a rationalizing relation, but merely at least as good as all elements $y \in S$ such that $(S, y) \notin \mathcal{N}$. That is, if the choice of y from S is already prohibited by the norm, there is no need that x dominates such an element y according to the rationalization. Needless to say, the chosen element x itself must be admissible in the presence of the prevailing system of external norms.

To make this concept of norm-conditional rationalizability precise, let a system of external norms $\mathcal{N} \in \mathbf{N}$ and a feasible set $S \in \Sigma$ be given. An $\mathcal{N}$-*admissible set* for $(\mathcal{N}, S)$, $A^{\mathcal{N}}(S) \subseteq S$, is defined by letting, for all $x \in S$,

$$x \in A^{\mathcal{N}}(S) \;\Leftrightarrow\; (S, x) \notin \mathcal{N}.$$

Note that, by assumption, $A^{\mathcal{N}}(S) \neq \emptyset$ for all $\mathcal{N} \in \mathbf{N}$ and for all $S \in \Sigma$.

4 Norm-Conditional Rationalizability

We now introduce our analytical framework and define the basic concepts of norm-conditional rationalizability. Let $R \subseteq X \times X$ be a (binary) relation on the non-empty universal set X. The asymmetric factor $P(R)$ of R is given by $(x, y) \in P(R)$ if and only if $(x, y) \in R$ and $(y, x) \notin R$ for all $x, y \in X$, and the symmetric factor $I(R)$ of R is defined by $(x, y) \in I(R)$ if and only if $(x, y) \in R$ and $(y, x) \in R$ for all $x, y \in X$.

The *transitive closure* $tc(R)$ of a relation R is defined by letting, for all $x, y \in X$,

$$(x, y) \in tc(R) \quad \Leftrightarrow \quad \text{there exist } K \in \mathbb{N} \text{ and } x^0, \ldots, x^K \in X \text{ such that}$$
$$x = x^0 \text{ and } (x^{k-1}, x^k) \in R \text{ for all } k \in \{1, \ldots, K\} \text{ and } x^K = y.$$

For any binary relation R, $tc(R)$ is the smallest transitive superset of R.

A relation $R \subseteq X \times X$ is *reflexive* if, for all $x \in X$,

$$(x, x) \in R$$

and R is *complete* if, for all $x, y \in X$ such that $x \neq y$,

$$(x, y) \in R \text{ or } (y, x) \in R.$$

R is *transitive* if, for all $x, y, z \in X$,

$$[(x, y) \in R \text{ and } (y, z) \in R] \;\Rightarrow\; (x, z) \in R.$$

It is clear that R is transitive if and only if $R = tc(R)$. A *quasi-ordering* is a reflexive and transitive relation and an *ordering* is a complete quasi-ordering.

R is *consistent* if, for all $x, y \in X$,

$$(x, y) \in tc(R) \;\Rightarrow\; (y, x) \notin P(R).$$

This notion of consistency is due to Suzumura (1976b) and it is equivalent to the requirement that any cycle must be such that all relations involved in this cycle are instances of indifference—strict preference cannot occur. To facilitate the understanding of this concept, we may define the *consistent closure* $cc(R)$ of R as the smallest consistent superset of R. This is the concept coined by Bossert, Sprumont and Suzumura (2005), which may be written explicitly as follows. For all $x, y \in X$,

$$(x, y) \in cc(R) \;\Leftrightarrow\; (x, y) \in R \text{ or } [(x, y) \in tc(R) \text{ and } (y, x) \in R].$$

Clearly, for any binary relation R, we have $R \subseteq cc(R) \subseteq tc(R)$ and R is consistent if and only if $R = cc(R)$. It is easy to verify that consistency implies (but is not implied by) the well-known *acyclicity* axiom which rules out the existence of strict preference cycles

(cycles composed entirely of instances of strict preference). Consistency and *quasi-transitivity,* which requires that $P(R)$ is transitive, are independent. Transitivity implies consistency but the reverse implication is not true in general. However, if R is reflexive and complete, consistency and transitivity are equivalent.

A relation R^* is an *extension* of R if and only if $R \subseteq R^*$ and $P(R) \subseteq P(R^*)$. If an extension R^* of R is an ordering, we refer to R^* as an *ordering extension* of R. One of the most fundamental results on extensions of binary relations is due to Szpilrajn (1930) who showed that any transitive and asymmetric relation has a transitive, asymmetric and complete extension. The result remains true if asymmetry is replaced with reflexivity, that is, *any quasi-ordering has an ordering extension.* Arrow (1951: 64) stated this generalization of Szpilrajn's theorem without a proof and Hansson (1968) provided a proof on the basis of Szpilrajn's original theorem. While the property of being a quasi-ordering is sufficient for the existence of an ordering extension of a relation, this is not necessary. As shown by Suzumura (1976b), consistency is *necessary and sufficient* for the existence of an ordering extension; see Suzumura (1976b: 389f).

We say that a choice function C on Σ is $\mathcal{N}$*-rationalizable* if and only if there exists a binary relation $R^{\mathcal{N}} \subseteq X \times X$ such that, for all $S \in \Sigma$ and for all $x \in S$,

$$x \in C(S) \;\Leftrightarrow\; x \in A^{\mathcal{N}}(S) \text{ and } [(x, y) \in R^{\mathcal{N}} \text{ for all } y \in A^{\mathcal{N}}(S)].$$

In this case, we say that $R^{\mathcal{N}}$ $\mathcal{N}$*-rationalizes* C, or $R^{\mathcal{N}}$ is an $\mathcal{N}$*-rationalization* of C.

To facilitate our analysis of $\mathcal{N}$-rationalizability, a generalization of the notion of the *direct revealed preference relation* $R_C \subseteq X \times X$ of a choice function C is of use. For all $x, y \in X$,

$$(x, y) \in R_C \;\Leftrightarrow\; \text{there exists } S \in \Sigma \text{ such that } [x \in C(S) \text{ and } y \in A^{\mathcal{N}}(S)].$$

The *(indirect) revealed preference relation* of C is the transitive closure $tc(R_C)$ of the direct revealed preference relation R_C.

We consider three basic versions of norm-conditional rationalizability. The first is $\mathcal{N}$-rationalizability by itself, where an $\mathcal{N}$-rationalization $R^{\mathcal{N}}$ does not have to possess any additional property (such as reflexivity, completeness, consistency or transitivity).

This notion of rationalizability is equivalent to $\mathcal{N}$-rationalizability by a reflexive relation (this is also true for the standard definition of rationalizability without external norms; see Richter 1971). The second is $\mathcal{N}$-rationalizability by a consistent relation (again, reflexivity can be added and an equivalent condition is obtained; see Bossert, Sprumont and Suzumura 2005). Finally, we consider $\mathcal{N}$-rationalizability by a transitive relation which, again as in the classical case, turns out to be equivalent to $\mathcal{N}$-rationalizability by an ordering; see Richter (1966; 1971).

We are now ready to identify a necessary and sufficient condition for each one of these notions of $\mathcal{N}$-rationalizability of a choice function. To obtain a necessary and sufficient condition for simple $\mathcal{N}$-rationalizability (that is, $\mathcal{N}$-rationalizability by a binary relation $R^{\mathcal{N}}$ that does not have to possess any further property), we follow Richter (1971) by generalizing the relevant axiom in his approach in order to accommodate an externally imposed system of norms $\mathcal{N}$. This leads us to the following axiom.

$\mathcal{N}$-conditional direct-revelation coherence: For all $S \in \Sigma$ and for all $x \in A^{\mathcal{N}}(S)$,

$[(x, y) \in R_C$ for all $y \in A^{\mathcal{N}}(S)] \;\Rightarrow\; x \in C(S)$.

Our first result establishes that this property is indeed necessary and sufficient for $\mathcal{N}$-rationalizability.

Theorem 1 *Let $\mathcal{N} \in \mathbf{N}$ be a system of external norms and let C be a choice function. C is $\mathcal{N}$-rationalizable if and only if C satisfies $\mathcal{N}$-conditional direct-revelation coherence.*

As is the case for the traditional model of rational choice on general domains, it is straightforward to verify that $\mathcal{N}$-rationalizability by a reflexive relation is equivalent to $\mathcal{N}$-rationalizability without any further properties of an $\mathcal{N}$-rationalization; this can be verified analogously to Richter (1971). However, adding completeness as a requirement leads to a stronger notion of $\mathcal{N}$-rationalizability; see again Richer (1971).

Next, we examine $\mathcal{N}$-rationalizability by a consistent relation, which is equivalent to $\mathcal{N}$-rationalizability by a reflexive and consistent relation. As in the traditional case, adding completeness, however, leads to a stronger property, namely, one that is equivalent to $\mathcal{N}$-rationalizability by an ordering; see Bossert, Sprumont and Suzumura (2005) for an analogous observation in the traditional model.

The requisite necessary and sufficient condition is obtained from $\mathcal{N}$-conditional direct-revelation coherence by replacing R_C with its consistent closure $cc(R_C)$.

$\mathcal{N}$-conditional consistent-closure coherence: For all $S \in \Sigma$ and for all $x \in A^{\mathcal{N}}(S)$,

$$[(x, y) \in cc(R_C) \text{ for all } y \in A^{\mathcal{N}}(S)] \;\Rightarrow\; x \in C(S).$$

We obtain

Theorem 2 *Let $\mathcal{N} \in \mathbf{N}$ be a system of external norms and let C be a choice function. C is $\mathcal{N}$-rationalizable by a consistent relation if and only if C satisfies $\mathcal{N}$-conditional consistent-closure coherence.*

Our final result establishes a necessary and sufficient condition for $\mathcal{N}$-rationalizability by a transitive relation which is equivalent to $\mathcal{N}$-rationalizability by an ordering. We leave it to the reader to verify that the proof strategy employed by Richter (1966; 1971) in the traditional case generalizes in a straightforward manner to the norm-dependent model when establishing that transitive $\mathcal{N}$-rationalizability is equivalent to $\mathcal{N}$-rationalizability by an ordering.

The requisite necessary and sufficient condition is obtained from $\mathcal{N}$-conditional direct-revelation coherence by replacing R_C with its transitive closure $tc(R_C)$.

$\mathcal{N}$-conditional transitive-closure coherence: For all $S \in \Sigma$ and for all $x \in A^{\mathcal{N}}(S)$,

$$[(x, y) \in tc(R_C) \text{ for all } y \in A^{\mathcal{N}}(S)] \;\Rightarrow\; x \in C(S).$$

We obtain

Theorem 3 *Let $\mathcal{N} \in \mathbf{N}$ be a system of external norms and let C be a choice function. C is $\mathcal{N}$-rationalizable by a transitive relation if and only if C satisfies $\mathcal{N}$-conditional transitive-closure coherence.*

5 Alternative Formulations

Our model of norm-conditional choice may appear somewhat restrictive at first sight because it specifies pairs of a feasible set and *a single object* not to be chosen from that set. One might want to consider the following seeming generalization of this approach: instead of only including pairs of the form (S, x) with $x \in S$ when defining a system of norms, one could include pairs such as (S, S') with $S' \subseteq S$, thus postulating that the subset S' should not be chosen from S. Contrary to first appearance, this does not really provide a more general model of norm-conditional rationalizability because, in order to formulate our notion of norm-conditional rationality, we require that a chosen element $x \in C(S)$ has to be at least as good as all feasible elements except those that are already excluded by the external norm according to a norm-conditional rationalization—that is, x has to be at least as good as all $y \in S$ except for those $y \in S$ such that $(S, y) \in \mathcal{N}$. Allowing for pairs (S, S') does not provide a more general notion of norm-conditional rationalizability because the subset of S, the elements of which have to be dominated by a chosen object, can be obtained in any arbitrary way from the subsets S' such that S' cannot be selected from S according to the external norm. For simplicity of exposition, we work with the simpler version of our model but note that this formulation does not involve any loss of generality when it comes to the definition of norm-conditional rationality employed in this paper.

An alternative way of modelling external norms is based on a suggestion of one of the referees. It employs a *norm-embodying* relation $\mathbf{R} \subseteq (X \times \mathcal{X}) \times (X \times \mathcal{X})$, where a pair (x, S) with $x \in S \subseteq X$ stands for the act of choosing x from a choice environment S given an underlying norm. The statement $((x, S), (y, T)) \in \mathbf{R}$ can then be interpreted to mean that choosing x from the admissible choice environment S is at least as good as choosing y from T in terms of conformity with the requisite external norm. The basic difference between such an approach and ours is that the former works with a *comparative* notion of norm satisfaction, whereas our model treats external norms as *constraints* on acceptable choice behavior. We do not view one setup as being superior to the other but we believe that our formulation in terms of constraints is more in the spirit of Sen's (1993) example because it allows us to rule out choice behavior that is not permitted in the sense of violating the requisite norm constraints. With an alternative model that is based on the relation 'is at least as norm-conforming as,' it seems to us that it is more difficult to capture the situations Sen had in mind.

This paper does not constitute the first attempt to accommodate Sen's (1993) criticism within a suitably modified theory of rational choice. In response to Sen (1993), Baigent and Gaertner (1996) employ a non-standard notion of rationalizability that obeys the restriction imposed by the external norm not to choose the *uniquely* greatest element according to some relation but behaves as the traditional version of rationalizability when the set of greatest objects contains at least two elements. A choice function that is rationalizable in their sense selects all second-greatest elements according to a rationalizing relation if there is a unique greatest element; if there are several greatest elements, C chooses *all* of these greatest elements. Baigent and Gaertner (1996: 241) state that they axiomatize the maxim "always choose the second largest except in those cases where there are at least two pieces which are largest, being of equal size. In that case, either may be chosen." However, this informal maxim appears to be in conflict with the formal definition and characterization provided by Baigent and Gaertner (1996: 243): if there is no unique greatest element, *all* greatest elements are chosen and not just one of them. Thus, there is a gap between their formal axiomatization and the informal maxim, the axiomatization of the latter being left unaccomplished so far. Moreover, their model is restricted to a rather narrow class of choice problems. If, instead of having two apples in the feasible set, the decision-maker faces a fruit basket containing one apple and one orange, picking the second-greatest element according to some rationalizing relation no longer seems to represent reasonable behavior: the fact that I as the decision-maker prefer apples to oranges, say, does not mean other people have the same preferences and, as a consequence, norms of politeness and decency do *not* dictate the choice of the orange—what if everyone else at the table prefers oranges to apples?

Gaertner and Xu (1999a) discuss an alternative approach covering cases where external norms may lead to the choice of the *median* alternative(s) according to some antisymmetric relation on X. This approach is compared to the traditional rational choice setup and to the Baigent and Gaertner (1996) framework in the antisymmetric case in Gaertner and Xu (1997) and in a more general setting in Gaertner and Xu (1999b).

An alternative type of norm-constrained choice is characterized in Gaertner and Xu (2004). The choice functions analyzed in this contribution have a domain that contains the empty set in addition to all non-empty subsets of X and, moreover, choice sets may be empty even if feasible sets are non-empty. The behavior Gaertner and Xu (2004) attempt to capture is the refusal to make a choice in response to the suppression of alternatives: if there is but a single alternative available, the decision-maker may choose the empty set as a means of expressing his or her displeasure with the suppression of other feasible alternatives. An example put forward by Sen (1997: 755) and used by Gaertner and Xu (2004) as a motivation of their approach is that of a government that outlaws all newspapers but one that is owned by the government itself. They argue that if several papers are available, the government paper may well be the choice of a decision-maker, but the absence of any alternative sources leads the agent to boycott the single available news outlet.

Xu (2007) discusses some special cases of norm-conditional rationalizability, namely, a variant of the Baigent and Gaertner (1996) 'never choose the uniquely largest' rule, the median-based rule (see Gaertner and Xu 1999a) and two versions of the 'protest-based' norm of Gaertner and Xu (2004). These special cases are obtained by ruling out the choice of unique best elements, elements better than the (bottom) median element, and non-empty choices in the case of single-valued feasible sets. See also

Baigent (2007) for a summary and discussion of the relevant literature.

All of the above responses to Sen's criticism are geared towards the characterization of a *specific* type of choice behavior and the notion of rationality that is used is, in turn, specific to the particular external norm under consideration. In contrast, our approach focuses on a general method to incorporate *any* external norm one might wish to specify.

6 Conclusion

Sen (1997) took an important step towards the norm-conditional theory of rationalizability through the concept of *self-imposed choice constraints,* excluding the choice of some alternatives from permissible conducts. Let $M(S, R)$ denote the set of R-maximal elements in S according to R, that is, the set of all elements of S that are not strictly preferred by any other element in S. According to Sen's (1997: 769) scenario, "the person may first restrict the choice options ... by taking a 'permissible' subset $K(S)$, reflecting *self-imposed* constraints, and then seek the maximal elements $M(K(S), R)$ in $K(S)$." Despite an apparent family resemblance between Sen's concept of self-imposed choice constraints and our concept of norm-conditionality, Sen did not go as far as to bridge the idea of norm-induced constraints and the theory of rationalizability as we did in this paper.

It is hoped that the present paper provides the missing link in the existing literature and shows that external norms can be made neatly compatible with a suitably modified revealed preference theory.

7 Proofs

We first provide three preliminary results. The following lemma states that the direct revealed preference relation R_C must be respected by any $\mathcal{N}$-rationalization $R^{\mathcal{N}}$. This observation parallels that of Samuelson (1948; 1950) in the traditional framework; see also Richter (1971).

Lemma 1 *Let $\mathcal{N} \in \mathbf{N}$ be a system of external norms and let C be a choice function. If $R^{\mathcal{N}}$ is an $\mathcal{N}$-rationalization of C, then $R_C \subseteq R^{\mathcal{N}}$.*

Proof. Suppose that $R^{\mathcal{N}}$ is an $\mathcal{N}$-rationalization of C and $x, y \in X$ are such that $(x, y) \in R_C$. By definition of R_C, there exists $S \in \Sigma$ such that $x \in C(S)$ and $y \in A^{\mathcal{N}}(S)$. Because $R^{\mathcal{N}}$ is an $\mathcal{N}$-rationalization of C, we obtain $(x, y) \in R^{\mathcal{N}}$. Thus, $R_C \subseteq R^{\mathcal{N}}$ must be true. ■

Analogously, any consistent $\mathcal{N}$-rationalization $R^{\mathcal{N}}$ must respect not only the direct revealed preference relation R_C but also its consistent closure $cc(R_C)$.

Lemma 2 *Let $\mathcal{N} \in \mathbf{N}$ be a system of external norms and let C be a choice function. If $R^{\mathcal{N}}$ is a consistent $\mathcal{N}$-rationalization of C, then $cc(R_C) \subseteq R^{\mathcal{N}}$.*

Proof. Suppose that $R^{\mathcal{N}}$ is a consistent $\mathcal{N}$-rationalization of C and $x, y \in X$ are such that $(x, y) \in cc(R_C)$. By definition of the consistent closure of a binary relation, $(x, y) \in R_C$ or $[(x, y) \in tc(R_C)$ and $(y, x) \in R_C]$ must hold. If $(x, y) \in R_C$, $(x, y) \in R^{\mathcal{N}}$ follows from **Lemma 1**. If $[(x, y) \in tc(R_C)$ and $(y, x) \in R_C]$, there exist $K \in \mathbb{N}$ and $x^0, \ldots, x^K \in X$ such that $x = x^0$, $(x^{k-1}, x^k) \in R_C$ for all $k \in \{1, \ldots, K\}$ and $x^K = y$. By **Lemma 1**, $(x^{k-1}, x^k) \in R^{\mathcal{N}}$ for all $k \in \{1, \ldots, K\}$ and, thus, $(x, y) \in tc(R^{\mathcal{N}})$. Furthermore, $(y, x) \in R_C$ implies $(y, x) \in R^{\mathcal{N}}$ by **Lemma 1** again. If $(x, y) \notin R^{\mathcal{N}}$, it follows that $(y, x) \in P(R^{\mathcal{N}})$ in view of $(y, x) \in R^{\mathcal{N}}$. Because $(x, y) \in tc(R^{\mathcal{N}})$, this contradicts the consistency of $R^{\mathcal{N}}$. Therefore, $(x, y) \in R^{\mathcal{N}}$. Thus, $cc(R_C) \subseteq R^{\mathcal{N}}$ must be true. ■

Finally, if transitivity is required as a property of an $\mathcal{N}$-rationalization $R^{\mathcal{N}}$, this relation must respect the transitive closure $tc(R_C)$ of R_C.

Lemma 3 *Let $\mathcal{N} \in \mathbf{N}$ be a system of external norms and let C be a choice function. If $R^{\mathcal{N}}$ is a transitive $\mathcal{N}$-rationalization of C, then $tc(R_C) \subseteq R^{\mathcal{N}}$.*

Proof. Suppose that $R^{\mathcal{N}}$ is a transitive $\mathcal{N}$-rationalization of C and $x, y \in X$ are such that $(x, y) \in tc(R_C)$. By definition of the transitive closure of a binary relation R_C, there exist $K \in \mathbb{N}$ and $x^0, \ldots, x^K \in X$ such that $x = x^0$, $(x^{k-1}, x^k) \in R_C$ for all $k \in \{1, \ldots, K\}$ and $x^K = y$. By **Lemma 1**, we obtain $x = x^0$, $(x^{k-1}, x^k) \in R^{\mathcal{N}}$ for all $k \in \{1, \ldots, K\}$ and $x^K = y$. Repeated application of the transitivity of $R^{\mathcal{N}}$ implies $(x, y) \in R^{\mathcal{N}}$. Thus $tc(R_C) \subseteq R^{\mathcal{N}}$ must hold. ■

Proof of Theorem 1. "Only if." Suppose $R^{\mathcal{N}}$ is an $\mathcal{N}$-rationalization of C. Let $S \in \Sigma$ and $x \in A^{\mathcal{N}}(S)$ be such that $(x, y) \in R_C$ for all $y \in A^{\mathcal{N}}(S)$. By **Lemma 1**, $(x, y) \in R^{\mathcal{N}}$ for all $y \in A^{\mathcal{N}}(S)$, which implies $x \in C(S)$ because $R^{\mathcal{N}}$ is an $\mathcal{N}$-rationalization of C.

"If." Suppose C satisfies $\mathcal{N}$-conditional direct-revelation coherence. We complete the proof by showing that $R^{\mathcal{N}} = R_C$ is an $\mathcal{N}$-rationalization of C. Let $S \in \Sigma$ and $x \in A^{\mathcal{N}}(S)$.

Suppose first that $x \in C(S)$. By definition, it follows immediately that $(x, y) \in R_C = R^{\mathcal{N}}$ for all $y \in A^{\mathcal{N}}(S)$.

Conversely, suppose that $(x, y) \in R_C = R^{\mathcal{N}}$ for all $y \in A^{\mathcal{N}}(S)$. It follows that $\mathcal{N}$-conditional direct-revelation coherence immediately implies $x \in C(S)$. Thus, C is $\mathcal{N}$-rationalizable by $R^{\mathcal{N}} = R_C$. ■

Proof of Theorem 2. The proof is analogous to that of **Theorem 1**. All that needs to be done is replace R_C with $cc(R_C)$ and invoke **Lemma 2** instead of **Lemma 1**. ■

Proof of Theorem 3. Again, the proof is analogous to that of **Theorem 1**. All that needs to be done is replace R_C with $tc(R_C)$ and invoke **Lemma 3** instead of **Lemma 1**. ■

References

Arrow, K.J. 1951; second ed. 1963. *Social choice and individual values.* New York: John Wiley & Sons.

Arrow, K.J. 1959. Rational choice functions and orderings. *Economica* 26: 121–7.

Baigent, N. 2007. Choice, norms and revealed preference. *Analyse & Kritik* 29: 139–45.

Baigent, N. and W. Gaertner. 1996. Never choose the uniquely largest: a characterization. *Economic Theory* 8: 239–49.

Bossert, W. 2001. Choices, consequences, and rationality. *Synthese* 129: 343–69.

Bossert, W., Y. Sprumont and K. Suzumura. 2005. Consistent rationalizability. *Economica* 72: 185–200.

Bossert, W., Y. Sprumont and K. Suzumura. 2006. Rationalizability of choice functions on general domains without full transitivity. *Social Choice and Welfare* 27: 435–58.

Gaertner, W. and Y. Xu. 1997. Optimization and external reference: a comparison of three axiomatic systems. *Economics Letters* 57: 57–62.

Gaertner, W. and Y. Xu. 1999a. On rationalizability of choice functions: a characterization of the median. *Social Choice and Welfare* 16: 629–38.

Gaertner, W. and Y. Xu. 1999b. On the structure of choice under different external references. *Economic Theory* 14: 609–20.

Gaertner, W. and Y. Xu. 2004. Procedural choice. *Economic Theory* 24: 335–49.

Hansson, B. 1968. Choice structures and preference relations. *Synthese* 18: 443–58.

Houthakker, H. 1950. Revealed preference and the utility function. *Economica* 17: 159–74.

Richter, M.K. 1966. Revealed preference theory. *Econometrica* 34: 635–45.

Richter, M.K. 1971. Rational choice. In *Preferences, utility, and demand,* eds. J. Chipman, L. Hurwicz, M. Richter and H. Sonnenschein, 29–58. New York: Harcourt Brace Jovanovich.

Samuelson, P.A. 1938. A note on the pure theory of consumer's behaviour. *Economica* 5: 61–71.

Samuelson, P.A. 1947. *Foundations of economic analysis.* Cambridge, Mass.: Harvard University Press.

Samuelson, P.A. 1948. Consumption theory in terms of revealed preference. *Economica* 15: 243–53.

Samuelson, P.A. 1950. The problem of integrability in utility theory. *Economica* 17: 355–85.

Sen, A.K. 1971. Choice functions and revealed preference. *Review of Economic Studies* 38: 307–17.

Sen, A.K. 1993. Internal consistency of choice. *Econometrica* 61: 495–521.

Sen, A.K. 1997. Maximization and the act of choice. *Econometrica* 65: 745–79.

Suzumura, K. 1976a. Rational choice and revealed preference. *Review of Economic Studies* 43: 149–58.

Suzumura, K. 1976b. Remarks on the theory of collective choice. *Economica* 43: 381–90.

Suzumura, K. 1977. Houthakker's axiom in the theory of rational choice. *Journal of Economic Theory* 14: 284–90.

Suzumura, K. 1983. *Rational choice, collective decisions and social welfare.* New York: Cambridge University Press.

Szpilrajn, E. 1930. Sur l'extension de l'ordre partiel. *Fundamenta Mathematicae* 16: 386–9.

Xu, Y. 2007. Norm-constrained choices. *Analyse & Kritik* 29: 329–39.