Revealed Conflicting Preferences: Rationalizing Choice with Multi-Self Models*

Attila Ambrus† Kareen Rozen‡

Harvard University Yale University

December 2008

(First Version: April 2008)

Abstract

We model a DM as a collection of utility functions (*selves*, *rationales*) and an aggregation rule (a theory of how selves are activated by choice sets) on which we impose five simple axioms of social choice. This framework encompasses many multi-self models proposed in the existing literature. For a broad class of aggregators we show that with sufficiently many selves the resulting model can rationalize *any* choice function. We define an accounting procedure for IIA violations and show that for any fixed number of selves, a lower bound on the set of choice functions that these aggregators can rationalize is given by the set of choice functions that exhibit no more IIA violations than a certain linear function of the number of selves.

JEL Codes: D11, D13, D71
Keywords: Multiple selves, IIA violations, rationalizability, complexity

---

*We are grateful to Eddie Dekel, Drew Fudenberg, John Geanakoplos, Dino Gerardi, Tzachi Gilboa, Jerry Green, Daniel Hojman, Gil Kalai, Bart Lipman, Philippe Mongin, Ben Polak, Ariel Rubinstein, Larry Samuelson, Rani Spiegler, and Tomasz Strzalecki for helpful suggestions. We also thank seminar audiences at Brown, Harvard, MIT, NYU, UCL and Yale for insightful comments.

†Department of Economics. Email: ambrus@fas.harvard.edu

‡Department of Economics and Cowles Foundation. Email: kareen.rozen@yale.edu

"There is very likely no unique method used by minds to make decisions. It is well known that individuals are generally not very logical, and that their decision behavior can be modified by the surrounding culture or by the acquisition of some special skill. In spite of this, it has to be admitted that, given a specific decision, a specific mind will use a specific method"
- Kenneth J. Arrow and Hervé Reynaud (1986)

1 Introduction

The classical model of choice endows the decision-maker (DM) with a single preference relation that she uses to select the best element from any set of alternatives. The single implication of this model is context-independent behavior, or the *Independence of Irrelevant Alternatives* (IIA), which dictates that if an alternative is deemed optimal in a set, it must remain optimal in any subset.[1] Consequently, a growing body of evidence suggesting that behavior is prone to context-dependence has spurred interest in alternative models of decision-making that can facilitate violations of IIA. In particular, since the seminal work of May (1954), many papers have proposed models of multi-self decision making to accommodate such behaviors.[2] This literature includes Kalai, Rubinstein and Spiegler (2002), Fudenberg and Levine (2006), Manzini and Mariotti (2007), and Green and Hojman (2007) in economics; Tversky (1969), Shafir, Simonson and Tversky (1993) and Tversky and Simonson (1993) in psychology; and Kivetz, Netzer and Srinivasan (2004) in marketing.[3] Often these models are motivated by the desire to explain a particular empirically observed choice behavior not consistent with rational choice. Some fix the number of selves (e.g., the dual-self model of Fudenberg and Levine (2006)) while others leave the number unrestricted (e.g., Kalai et al. (2002)).

[1]This also implies transitive choice behavior, which is often violated in experimental settings (e.g., see Tversky (1969) and Lee, Amir and Ariely (2007)).

[2]Another approach, developed in Bernheim and Rangel (2007) and Salant and Rubinstein (2008), allows for context-dependence by considering extended choice situations where behavior can depend on unspecified *ancillary conditions* or *frames*. While information effects can explain some context dependence (Sen (1993), Kochov (2007), Kamenica (forthcoming)), they cannot explain many systematic violations of IIA (Tversky and Simonson (1993)).

[3]An expanded shortlist of the multiple-selves or multiple-utility literature includes Benabou and Pycia (2002), Masatlioglu and Ok (2005), Evren and Ok (2007), and Chatterjee and Krishna (forthcoming). This literature is also related to the application of social choice tools in multi-criteria decision problems, as in Arrow and Raynaud (1986), and is related more generally to the theory of multiattribute utility (see Keeney and Raiffa (1993)).

There has been little effort to connect the various models, or to conduct an analysis of multi-self decision-making using a more systematic approach. In this paper, we develop a framework to examine a DM with multiple selves, when choice sets themselves serve as frames that influence how the preferences of different selves get aggregated. More formally, we propose to model the DM as a collection of utility functions U (selves or rationales) and an *aggregation rule* f (decision-making method) that combines these utility functions in a possibly context-dependent way. That is, given a choice set A, and selves U, aggregator f specifies an aggregated utility for every alternative in A. An aggregator corresponds to a theory of how selves are activated by choice sets. We posit only that the aggregator satisfies five simple axioms from social choice theory. The multi-self models proposed in the literature can be translated into this framework, and many of them satisfy the axioms we impose on the aggregation rule.

Our main point of interest is investigating the set of behaviors that a specific model of multi-self decision-making, as captured by a given aggregation rule f, can rationalize (explain). We address this question both with a fixed number of selves, as well as with no *a priori* restriction. Formally, we assume that the DM's behavior is described by a choice function c, which specifies the alternative she selects in each subset of some grand set of alternatives X. We say that a DM's choice function is *rationalized* by a finite collection of selves U and an aggregator f if the choice function selects the unique maximizer of aggregate utility $f \circ U$ in every choice set. For some aggregators, it is straightforward to determine the set of choice functions that can be rationalized. For example, if the DM's method of aggregating the utilities of her various selves is simple utilitarianism, then the set of choice functions is exactly the set of rational choice functions, regardless of the number of selves. But what if the aggregator is the "normalized contextual concavity model" proposed in Kivetz et al. (2004),

$$\sum_{i=1}^{n}(\max_{a'\in A}\ u_i(a') - \min_{a'\in A}\ u_i(a')) \cdot \left[\frac{u_i(a) - \min_{a'\in A}\ u_i(a')}{\max_{a'\in A}\ u_i(a') - \min_{a'\in A}\ u_i(a')}\right]^{c_i}?$$

Our main result establishes that for a large class of aggregators, including various aggregators proposed in previous papers (and the example above), if there is no restriction on the number of selves, the model can rationalize any choice function. Hence, without knowledge of the number of motivations, the model has no testable restrictions on behavior. For the same set of aggregators, we provide a lower bound on the set of choice functions that can be rationalized given a fixed number of selves. In particular, we show that the model can rationalize any choice function that exhibits no more IIA violations than a certain linear function of the number of selves, where the number of IIA violations is defined by a simple accounting procedure.

The question of what range of behaviors a given aggregator can rationalize with a given number of selves can be asked in an alternative way, as what is the minimum number of selves required to rationalize a given choice function, with a given aggregator. Hence, our results also address the required *complexity* of a rationalization and connect it to the extent to which the choice behavior in question deviates from rationality, as measured by the number of IIA violations.[4]

Our main interpretation of the framework is one of individual decision-making with multiple motives. Psychologists have long viewed the *multiplicity of self* as a normal feature instead of a sign of pathology; and even psychologists who prefer a unitary view of the self accept that "the singular self is a hypothetical construct, an umbrella under which experiences are organized along various dimensions or motivational systems" and which "is fluid in that it shifts in different contexts as various motivations are activated" (Lachmann (1996)). This interpretation, namely that the decision-maker has multiple goals and resolves trade-offs among these in a manner affected by the choice set, fits our model as well as the more literal interpretation of multiple selves.

In line with the notion of aggregation in our model, psychologists believe that a fluid form of compromise among selves is necessary for healthy behavior.[5] The possibility of compromise is an important sense in which our model differs from Kalai et al. (2002) (henceforth KRS), who were the first to address whether a given choice behavior can be rationalized and to examine the complexity of the required rationalization. KRS propose that a collection of strict preference relations rationalizes a choice function if the choice from each set is optimal for at least one of the preference relations; they show that if there are n alternatives, then any choice behavior can be rationalized with $n-1$ rationales. In this view, each self serves as

[4]Complexity according to this approach is measured by the number of selves, which is analogous to measuring the complexity of finite automata by the number of states (e.g., see Salant (2007) in the decision-theoretic literature).

[5]This is as opposed to disassociated selves (i.e., overly autonomous selves), or a high self-concept differentiation (a lack of interrelatedness of selves across contexts) both of which are connected to pathological or unhealthy behavior; see Power (2007), Donahue, Robins, Roberts and John (1993), and Mitchell (1993).

a dictator for some subset of choices. In contrast, in our framework it can happen that the choice is not the most preferred alternative of any of the selves, but the *best compromise*, in the sense that it maximizes aggregate utility.

There are several recent contributions to the literature on multi-self decision-making which mostly focus on a different set of questions than we do. Of these, the most related is Green and Hojman (2007) (henceforth GH), who also explain choice behavior using certain structured aggregation methods. They consider scoring rules, a parametric family of ordinal voting rules able to rationalize any social choice behavior if one allows for any probability measure on strict preference orderings - a result shown in both Saari (1999) and GH. Scoring rules are ordinal and do not fall within our class of aggregation methods; furthermore, GH do not address the question of rationalizing choice functions using a restricted domain of selves, instead focusing attention on welfare analysis using sets of possible rationalizations.[6] Other related work includes Manzini and Mariotti (2007) and Cherepanov, Feddersen and Sandroni (2008), who consider sequential application of multiple rationales to eliminate alternatives, a process they show can rationalize certain choice functions. Finally, Fudenberg and Levine (2006) consider a dual-self model of dynamic choice, where the two selves' utilities are aggregated in a menu-dependent way.[7]

Besides the primary interpretation using multi-self individual decision-making, our results can also be used to analyze collective household choice. For this reason, we extend the analysis to incomplete choice functions, such as demand functions. Our results complement those of Browning and Chiappori (1998) and Chiappori and Ekeland (2006) in this context. We can also address questions regarding the size of the *subjective state-space* in models of choice over menus, complementing the results of Dekel, Lipman and Rustichini (2001).

This paper is organized as follows. Section 2 presents the framework and examples thereof. Section 2.4 shows how some rule-of-thumb decision-procedures can be rationalized within this framework. Section 3 describes our accounting procedure for IIA violations and Section 4 presents our main results. Section 5 considers two applications of our model. Finally, Section 6 investigates different ways of tightening the bounds on rationalization.

[6]Bernheim and Rangel (2007) also focus on welfare analysis given choices contradicting rational decision-making.

[7]See also Chatterjee and Krishna (forthcoming) for a model of dual-self decision-making.

2 A framework for rationalizing choice

2.1 Main concepts and definitions

Suppose that we observe a DM's choice behavior on a finite set of alternatives X. Denote by $P(X)$ the set of nonempty subsets of X. The DM's *choice function* $c : P(X) \to X$ identifies the alternative $c(A) \in A$ that she chooses from each $A \in P(X)$. A *rationalization* of the DM's choice function consists of two components, a collection of *selves* U and an *aggregator* f that combines these. The DM's selves represent her conflicting motivations or priorities. The aggregator corresponds to the DM's method of "sorting out" her priorities to come to a decision.

Formally, given a basic set of alternatives X, a self (a.k.a. *reason*, *rationale*) is a utility function $u : X \to \mathbb{R}$. Hence, each self is an element of the function space $\mathbb{R}^X$, and $u(x)$ is the utility level that self u allocates to $x \in X$. For each positive integer n, we denote by $\mathcal{U}^n(X) = \times_{i=1}^n \mathbb{R}^X$ the set of all n-tuples of selves defined over X, and by $\mathcal{U}(X) = \cup_{n=1}^\infty \mathcal{U}^n(X)$ the set of all finite tuples of selves over X. We will denote a particular collection of selves by U. To denote the number of selves in U, we use the notation $|U|$ or simply n when no confusion would arise.

An aggregator f specifies an aggregate utility for every alternative a in every choice set A, given any (finite) basic set of alternatives X and any collection of selves U defined over these alternatives. Formally, the domain over which f is defined is $\{a, A, X, U | X \in \mathcal{X}, U \in \mathcal{U}(X), A \in P(X), a \in A\}$, where $\mathcal{X}$ is the set of all conceivable basic sets of alternatives, and every $X \in \mathcal{X}$ is finite. We indicate X explicitly among the arguments of f because we are interested in investigating how the number of selves needed to rationalize a given choice rule depends on the number of alternatives in X. Note that since the choice set is one of the arguments of the function, f aggregates the utilities of the selves in a possibly context-dependent way.[8]

Definition 2.1. *We say that a choice function $c(\cdot)$ on X is rationalized by the aggregator f if there exists a finite collection of selves $U \in \mathcal{U}(X)$ such that for every $A \in P(X)$, $c(A) = \arg\max_{a \in A} f(a, A, X, U)$.*

Although aggregation in the above framework is cardinal (intensities of preferences might matter), the model has the ordinal feature that there can be many

[8] We could also permit aggregators with restricted domains: let $\hat{\mathbb{R}}^X$ be a convex subset of $\mathbb{R}^X$ and define instead $\mathcal{U}^n = \times_{i=1}^n \hat{\mathbb{R}}^X$.

"equivalent" representations of an aggregator in this context. In particular, if f rationalizes the choice function c using the selves U, then so does any increasing transformation of f; and similarly, if f rationalizes c using the selves U, then $f \circ h^{-1}$ rationalizes c using the selves $h \circ U$, where $h : \mathbb{R} \to \mathbb{R}$ is invertible on the appropriate domain.

2.2 Basic axioms of aggregation

We are interested in examining theories of aggregation that are in line with the underlying selves' preferences. For this reason, for the rest of the paper we restrict attention to aggregators satisfying the following properties, most of which are familiar from the theory of social choice. As we will argue below, imposing these properties is a natural requirement if the aggregation of utilities is cardinal and the framing effect of a choice set operates only through the utility levels of alternatives for different selves.

To state these properties, we let $\pi : X \to X$ be any permutation and define $u \in \mathbb{R}^X$ to be δ-*indifferent* if $|u(a) - u(b)| < \delta$ for all $a, b \in X$. We define $U = (u_1, u_2, \ldots, u_n) \in \mathcal{U}(X)$ to be δ-indifferent if u_i is δ-indifferent for every i. Also, for any $U, U' \in \mathcal{U}(X)$, (U, U') denotes $(u_1, u_2, \ldots, u_{|U|}, u'_1, u'_2, \ldots, u'_{|U'|}) \in \mathcal{U}(X)$.

P1 (Neutrality) For any permutation π, $f(\pi(a), \pi(A), X, U \circ \pi^{-1}) = f(a, A, X, U)$.

P2 (Single-self respect) For any $u \in \mathbb{R}^X$, $u(a) \geq u(b)$ if and only if $f(a, A, X, u) \geq f(b, A, X, u)$.

P3 (Separability) If $f(a, A, X, U) \geq f(b, A, X, U)$ and $f(a, A, X, \hat{U}) \geq f(b, A, X, \hat{U})$ then $f(a, A, X, (U, \hat{U})) \geq f(b, A, X, (U, \hat{U}))$, with strict inequality if one of the above holds strictly.

P4 (Continuity at indifferent selves) If $f(a, A, X, U) > f(b, A, X, U)$ then for any $k \in Z_+$ there is $\delta_k > 0$ such that $f(a, A, X, (U, U')) > f(b, A, X, (U, U'))$ for any δ_k-indifferent $U' \in \mathcal{U}^k(X)$.

P5 (Duplication) If $U(a) = U(\hat{a})$ then $f(\cdot, A \cup \{a\}, X, U) = f(\cdot, A \cup \{\hat{a}\}, X, U)$.

Neutrality implies that the particular names of elements do not affect their ranking. Single-self respect is a minimal consistency requirement. Separability requires that if two collections of selves each prefer the alternative a to the alternative b, then these selves combined also prefer a to b. We note that Single-self respect and Separability together imply Pareto-optimality. Continuity at indifferent selves requires strict preference orderings implied by the aggregator to be robust to the addition of nearly-indifferent collections of selves. This is the axiom that separates the class of aggregators we study from ordinal ones, since repeated application of the axiom implies that one self's strict preference ordering is not reversed by arbitrarily many finite number of other selves, provided that the latter selves are all close enough to be indifferent (which only makes sense in a cardinal setting). Finally, Duplication says that aggregation is only affected by the utility levels of the alternatives in a given choice set. In particular, choice is not affected by which of two alternatives is adjoined to a set as long as those two alternatives yield exactly the same utility to all of the selves.

2.3 Examples of aggregators

The following are examples of context-dependent aggregators satisfying P1-P5, that are equivalent or closely related to models proposed in the existing literature.

Example 2.2 (Passion-driven and passion-muted models)**.** Suppose there is a strictly monotonic and continuous weighting function $g : R \to R$ such that for all $U \in \mathcal{U}$ and choice sets $A \subseteq X$,

$$f(a, A, X, U) = \sum_{i=1}^{n} g\big(\max_{b \in A} u_i(b) - \min_{b \in A} u_i(b)\big)\ u_i(a)$$

If $g(\cdot)$ is increasing, the model is a *passion-driven* one in which selves who are more "passionate" about the alternatives in the set A receive greater weight in the decision-process because they are more vociferous than selves who are more or less indifferent among the possibilities. If $g(\cdot)$ is decreasing,[9] the model may be seen as a *passion-muted* model or a context-dependent version of the models of relative utilitarianism in Dhillon and Mertens (1999) and Segal (2000), where a DM's weight in society is normalized by her utility range over the grand set of alternatives. Observe that a is preferred to b in the pair $\{a, b\}$ if and only if

$$\sum_{i=1}^{n} \underbrace{g(|u_i(a) - u_i(b)|)(u_i(a) - u_i(b))}_{\text{odd function of } u_i(a)-u_i(b)} > 0$$

[9]If $g(\cdot)$ is decreasing, we impose the restriction $\lim_{x \to 0^+} xg(x) = 0$.

Therefore, for pairwise choices the aggregator is similar to the *additive difference* model of Tversky (1969), which accounts for potentially intransitive pairwise choice behavior by positing utilities $v_1, v_2, \ldots, v_n$ and an odd function $\phi : \mathbb{R} \to \mathbb{R}$ such that $x \succ y$ if and only if $\sum_{i=1}^{n} \phi(v_i(x_i) - v_i(y_i)) > 0$. For larger choice sets, the aggregator can be thought of as a generalization of the additive difference model that permits context-dependence.

Example 2.3 (Loss aversion)**.** Suppose that the aggregator is given by

$$\sum_{i=1}^{n} \Big[(u_i(a) - \text{midpt}_i(A)) \cdot 1_{u_i(a) \geq \text{midpt}_i(A)} + \lambda_i (u_i(a) - \text{midpt}_i(A)) \cdot 1_{u_i(a) < \text{midpt}_i(A)} \Big],$$

where $\text{midpt}_i(A)$ is the midpoint of the range of u_i on A and λ_i is the loss aversion parameter for self i. This aggregator is a specific formulation of the model proposed in Tversky and Kahneman (1991), when the reference point for a self is the midpoint of the utility range given the choice set.

Example 2.4 (Costly self-control aggregators)**.** Fudenberg and Levine (2006) propose a dual-self impulse control model with a long-run self exerting costly self-control over a short-run self. The reduced-form model they derive has an analogous representation in our framework, with two selves: the long-run self, with utility given by u_{RF} (the expected present value of the utility stream induced by the choice in the present), and the short-run self, with utility function u (the present period consumption utility).[10] Using our terminology, the reduced form representation of their model assigns to alternative a the aggregate utility $u_{RF}(a) - C(a)$, where term $C(a)$ depends on the attainable utility levels for the short-run self and is labeled as the cost of self-control. For example, using Fudenberg and Levine (2006)'s parametrization, $C(a) = \gamma[\max_{a' \in A} u(a') - u(a)]^{\psi}$.

One way to generalize this aggregator to any number of selves would be to introduce multiple types of short-term temptations, represented by selves $u_2, ..., u_n$, and to define the aggregator

$$f(a, A, X, U) = u_1(a) - \sum_{i=2}^{n} \gamma[\max_{a' \in A} u_i(a') - u_i(a)]^{\psi}, \quad \text{where } u_1 = u_{RF}.$$

[10] The long-run self's utility is equal to the short-run self's utility plus the expected continuation value induced by the choice. If the latter can take any value, then u_{RF} is not restricted by the short-run utility u. If continuation values cannot be arbitrary (for example they have to be nonnegative) then u restricts the possible values of u_{RF}, hence U has a restricted domain. In Fudenberg and Levine (2006) the utility functions also depend on a state variable y. Here we suppress this variable, instead make the choice set explicit.

This is an example of an aggregator in our framework in which different selves are treated asymmetrically (here the long-run self is treated differently than the rest).

Example 2.5 (Contextual concavity models from marketing)**.** Kivetz et al. (2004) (henceforth KNS) considers various models capturing the compromise effect documented in experimental settings. KNS consider goods (e.g., laptops) which have defined attribute levels (e.g., processor speed) and posit utility levels ("partworths") for a given attribute. That is, they consider multiattribute alternatives and predefine the number of "selves" according to their selected good attributes. One type of model considered in KNS is referred to as a contextual concavity model. Using our notation, the simple contextual concavity model they propose is given by

$$f(a, A, X, U) = \sum_{i=1}^{n} (u_i(a) - \min_{a' \in A} \ u_i(a'))^{c_i},$$

where c_i is the concavity parameter and each i corresponds to the i-th attribute. They also propose a version that is normalized by the range of utilities, which we featured in Section 1.

In the examples above, the aggregator depends only on utility levels that are attainable in the choice set. One might also be interested in aggregators that give greatest weight to selves unhappy with the choice set (e.g., their average utility over the set is lower than their average utility from other menus). Our framework permits such dependence on unattainable utility levels, as demonstrated by the aggregator used in the following section to rationalize a natural decision-rule.

2.4 An example for rationalizing a decision rule

The *median procedure* is a simple choice rule defined in KRS. There is a strict ordering $\succ$ defined over elements of X, and the DM always chooses the median element of each $A \subseteq X$ according to $\succ$ (choosing the right-hand side element among the medians from choice sets with even number of alternatives).

To rationalize this behavior, we consider the following aggregator.

$$f(a, A, X, U) = \prod_{i=1}^{n} (u_i(a) + \max_{a' \in X} \ u_i(a') - \operatorname*{med}_{a' \in A} \ u_i(a')),$$

where $\operatorname*{med}_{a' \in A} u_i(a')$ is the median element of the set $\{u_i(a')\}_{a' \in A}$, with the convention that in sets with an even number of distinct utility levels, for odd i it is the higher one among the two median utility levels and for even i it is the lower one. The geometric aggregation implies that in case of selves having exactly the opposite preferences, the aggregated utility of an alternative from a given choice set is maximized when it is closest to the median element of the utility levels from the choice set.

Indeed, we claim that with the above aggregator, two selves can be used to rationalize the median procedure. Let $a_1, a_2, ..., a_N$ stand for the increasing ordering of alternatives in X according to $\succ$, and define $u_1(a_i) = i$ and $u_2(a_i) = N + 1 - i$ for all $i \in \{1, ..., N\}$. It is easy to see that it is indeed the median element of any choice set that maximizes f, since the sum of $u_1(a) + \max_{a' \in X} u_1(a') - \operatorname{med}_{a' \in A} u_1(a')$ and $u_2(a) + \max_{a' \in X} u_2(a') - \operatorname{med}_{a' \in A} u_2(a')$ is constant across all elements of X, and the two terms are equal at the median. Therefore the product that defines f is maximized at the median.

This rationalization is relatively simple and intuitive: the DM is torn between two motivations, one in line with ordering $\succ$, and one going in exactly the opposite direction. Moreover, the geometric aggregation of these preferences drives the DM to choose the most central element of any choice set. In contrast, KRS show that in their framework, in which exactly one self is responsible for any decision, as the size of X increases, the number of selves required to rationalize the median procedure goes to infinity.[11] While dictator-type aggregators as in KRS do not provide an intuitively appealing explanation for the median procedure, an aggregator that captures compromise along selves along the lines of the above-defined f does yield a model that rationalizes the median procedure in a simple and intuitive way.

There are many variants of the above aggregator that do not select exactly the median from every choice set, but have a tendency to induce the choice of a centrally located element from any choice set. For example, consider $f(a, A, X, U) = \prod_{i=1}^{n} (u_i(a) - \min_{a' \in X} u_i(a'))$. This aggregator allocates 0 aggregate utility to any element that is minimal for some self given the choice set, and strictly positive utility to all other elements (the largest one to the element maximizing the product of utility surpluses for different selves). In general, if f is menu-dependent and aggregates

[11] Another simple procedure considered in KRS which in their framework requires a large number of selves to rationalize is the second-best procedure, suggested first by Sen (1993). It is again possible to provide an aggregator fitting our framework such that two selves with opposite interests rationalize the choice rule - please contact the authors for details.

the utilities of selves through a concave function, the choice induced by f exhibits a *compromise effect* or *extremeness aversion*, as in the experiments of Simonson (1989): given two opposing motivations, an alternative is more likely to be selected the more centrally it is located. If, on the other hand, f is menu-dependent and convex, then it can give rise to a *polarization effect*, as in the experiments of Simonson and Tversky (1992): the induced choice is likely to be in one of the extremes of the choice set. Hence, our model can be used to reinterpret experimental choice data in different contexts, in terms of properties of the aggregator function.

3 Counting IIA violations

The examples of decision-rules presented in the previous section violate the Independence of Irrelevant Alternatives (IIA) because they are context-dependent.[12] IIA requires that if $a \in A \subset B$ and $c(B) = a$ then $c(A) = a$. This says that if an alternative is chosen from a set, then it should be chosen from any subset in which it is contained. It is well known that a choice function can be rationalized as the maximization of a single preference relation if and only if it has no violations of IIA. In the next section we connect the choice functions that a given aggregator can rationalize with a fixed number of selves to the number of IIA violations that a choice function exhibits. For this reason, below we formally define an accounting procedure for the number of IIA violations.

The number of IIA violations can be determined straightforwardly for choice functions over three-element sets; e.g., if the choice over pairs is transitive but the second-best element according to the pairs is selected from the triple, there is one violation of IIA. For a larger set of alternatives, there are different plausible ways to define the number of violations. For example, suppose that

$$
\begin{aligned}
c(\{a,b,c,d,e,f\}) &= d \\
c(\{a,b,c,d,e\}) &= b \\
c(\{a,b,c,d\}) &= b \\
c(\{b,c,d\}) &= c.
\end{aligned}
$$

In light of $c(\{a,b,c,d,e,f\}) = d$, IIA dictates that the last three choices should be

[12] Under the restriction of single-valued choice, the IIA condition is equivalent to Sen's α - see Sen (1971) - or WARP, the weak axiom of revealed preference.

d (but they are not). In light of $c(\{a,b,c,d,e\}) = b$, IIA dictates that the choice from $\{b,c,d\}$ should be b (but it is not), and the IIA implication for $\{b,c,d\}$ is again violated in light of $c(\{a,b,c,d\}) = b$. Hence, one way of counting would indicate five IIA violations with respect to the above four choice sets.

According to the above counting method, a given choice can cause many IIA violations. Instead, according to our counting procedure, any choice can increase the number of violations by at most one, and in the above example only the choices from $\{a,b,c,d,e\}$ and $\{b,c,d\}$ are associated with violations. The reason is that while $c(\{a,b,c,d\}) = b$ does contradict $c(\{a,b,c,d,e,f\}) = d$, the intermediate choice $c(\{a,b,c,d,e\}) = b$ itself implies by IIA that $c(\{a,b,c,d\}) = b$. In sum, our accounting procedure considers only the *first* violation of a choice, not further violations of the same choice in subsets of the set associated with the first violation.

Definition 3.1 (IIA violation)**.** *The set A causes an IIA violation under the choice function $c(\cdot)$ if (1) there exists B such that $A \subset B$ and $c(B) \in A \setminus \{c(A)\}$, and (2) for every A' such that $A \subset A' \subset B$, $c(A') \notin A$.*

Then, the total number of IIA violations is defined in the natural way.

Definition 3.2 (Number of IIA violations)**.** *The total number of IIA violations of a choice function $c(\cdot)$ is given by $\mathrm{IIA}(c) = \#\{A \in P(X) \mid A \text{ causes an IIA violation}\}$.*

We remark that one possible alternative measure of the number of IIA violations is the minimal number of sets at which the choice function would have to be changed to make it rational. This measure can in general be either larger or smaller than our measure of the number of IIA violations.[13]

4 Main results

We now present our main results, which give lower bounds on the choice functions that an aggregator can rationalize with a given number of selves. For ease of

[13]Indeed, suppose that pairwise choices exhibit the transitive ranking a preferred to b preferred to c. Under our measure, there is one violation of IIA if $c(\{a,b,c\}) = b$, which is defeated once in the pair $\{b,c\}$, and two violations of IIA if $c(\{a,b,c\}) = c$, which is defeated twice. The alternative measure counts one violation either way. To see that the alternative measure can also be larger, consider the choice function over $\{a,b,c,d,e\}$ which chooses the alphabetically-lowest alternative in all sets, except that b is chosen in three-element sets in which it is contained as well as from the pair $\{a,b\}$. The alternative measure counts four violations, while ours counts three. We thank both John Geanakoplos and Bart Lipman for suggesting this measure to us.

exposition, in this section we restrict attention to aggregators that only depend on alternatives in the choice set.

P6 (Independence of unavailable alternatives) For any basic sets of alternatives $X, X' \in \mathcal{X}$ such that $A \in X \cap X'$, and for any selves $U^X \in \mathcal{U}(X)$ and $U^{X'} \in \mathcal{U}(X')$ that agree on A (i.e., $U^{X'}(a) = U^X(a)$ for all $a \in A$), the aggregator satisfies $f(\cdot, A, X, U^X) = f(\cdot, A, X', U^{X'})$.

In Appendix B we extend our results to aggregators violating P6.

We start in Section 4.1 by demonstrating how to construct selves that rationalize a choice function in the case of the passion-driven aggregator. The construction provides intuition for the connection between the number of selves and the number of IIA violations. In Section 4.2 we generalize the construction to any aggregator satisfying a property that we call triple-solvability. This property holds for the aggregators in all of the examples we considered. In Section 4.3, we provide sufficient conditions for triple-solvability within the class of anonymous, additively separable and scale invariant aggregators. In particular, we show that triple solvability is broadly satisfied. For example, it is satisfied if the aggregator can rationalize "third-place choice," using the terminology of GH; or more generally if the aggregator "stretches" utility differences in a nonlinear way that we formalize below.

4.1 Rationalizing choice with passion-driven aggregation

Suppose that we are interested in rationalizing some choice function $c(\cdot)$ using the passion-driven aggregator, which is given by

$$f(a, A, X, U) = \sum_{i=1}^{n} g\big(\max_{b \in A} u_i(b) - \min_{b \in A} u_i(b)\big)\ u_i(a)$$

where $g(\cdot)$ is increasing. Before considering an arbitrary grand set of alternatives X, let us first examine how this aggregator behaves on an arbitrary three-element set of alternatives $\hat{X} = \{a, b, c\}$. Supposing that we were to use f to aggregate the five selves $U = (u_1, u_2, u_3, u_4, u_5)$ specified below, how would f evaluate each alternative in each subset of $\hat{X}$?[14]

u_1		u_2		u_3		u_4		u_5	
b	2	b	2	c	2	a, c	2	a	2
c	1	a	1	b	1	b	0	b, c	0
a	0	c	0	a	0				

It is easy to see that the aggregator selects a from the choice set $\{a, b\}$. Observe that $f(a, \{a, b\}, \hat{X}, U) = 4g(2) + g(1)$ and $f(b, \{a, b\}, \hat{X}, U) = 2g(2) + 3g(1)$, hence $f(a, \{a, b\}, \hat{X}, U) > f(b, \{a, b\}, \hat{X}, U)$ if and only if $g(2) > g(1)$, which holds since $g(\cdot)$ is strictly increasing. By contrast, the aggregator assigns equal utility to all alternatives in any other menu:

$$f(a, \{a, c\}, \hat{X}, U) = f(c, \{a, c\}, \hat{X}, U) = 2g(0) + g(1) + 2g(2)$$
$$f(b, \{b, c\}, \hat{X}, U) = f(c, \{b, c\}, \hat{X}, U) = 3g(1) + 2g(2)$$
$$f(a, \{a, b, c\}, \hat{X}, U) = f(b, \{a, b, c\}, \hat{X}, U) = f(c, \{a, b, c\}, \hat{X}, U) = 5g(2)$$

That is, a beats b when the choice set is $\{a, b\}$, while the selves cancel each other out for any other subset of $\hat{X}$. We call such a collection of selves defined on $\hat{X}$ a *triple-basis* for this aggregator. In the case of this aggregator, the selves above would still be a triple-basis if we were to scale all the utilities by a common constant.

Given an arbitrary X and any choice function c defined on X, we can use the triple-basis above to construct a collection of selves that rationalize c using the passion-driven aggregator f. The procedure works as follows. We examine all possible choice sets in X from smallest to largest, first going through all choice sets of size two, then all choice sets of size three, etc. We ignore any choice set that does not cause an IIA violation. For each choice set A that does cause an IIA violation, the construction creates a collection of selves U^A defined on X such that

1. $c(A)$ is selected under $f \circ U^A$ from every subset of A in which it is contained
2. The selves U^A cancel each other out under f on every other choice set (that is, on sets not containing $c(A)$ or sets containing some element of $X \setminus A$).
3. The selves U^A are "indifferent enough" so that their trickle-down effect does not overturn the strict preference of previously constructed selves

[14]In the i-th column, the alternative on the left is assigned the utility number to its right.

Finally, the construction creates an extra self u^*, that is indifferent enough never to overturn any of the other selves' strict preferences, in the standard way: the self allocates the highest utility to $c(X)$, the next highest utility to $X \setminus \{c(X)\}$, and so on. All in all, this procedure constructs a collection of $1 + 5 \cdot \text{IIA}(c)$ selves.

Using the triple-basis above, it is easy to construct the collection of selves U^A associated with a set A that causes an IIA violation. To satisfy the first two properties above, we simply let $c(A)$ play the role of a in the triple-basis, all the elements of $A \setminus \{c(A)\}$ play the role of b, and all the elements of $X \setminus A$ play the role of c. That is, we extend the utilities from $\{a, b, c\}$ to the given X such that: each self allocates the same utility to $c(A)$ as to a in the triple-basis, the same utility to elements of $A/c(A)$ as to b in the triple base, and the same utility to X/A as to c in the triple-basis. Neutrality (P1) and duplication (P5) then imply that the properties of the triple-basis carry over: for each $B \subseteq A$ that contains $c(A)$, $f(c(A), B, X, U^A) > f(y, B, X, U^A)$ for all $y \in B \setminus \{c(A)\}$, and for all other subsets $B' \subseteq X$, $f(x, B'^A) = f(y, B'^A)$ for all $x, y \in B'$. To satisfy the third property above, we can use continuity (P4) and scale all the selves in the triple-basis by some appropriately chosen $\varepsilon > 0$.

This entire collection of selves rationalizes $c(\cdot)$ under f. The construction ensures that $c(A)$ is selected from any set causing an IIA violation; one need only check that constructed selves do not interfere with choices associated with sets that do not cause IIA violations. To loosely illustrate the idea, consider any nested sequence of choice sets that decreases by one alternative. Given X, or any set from which c does not contradict the choice from X, all selves besides u^* are indifferent, hence by single-self respect (P2) and separability (P3) the preferences of u^* prevail. For the first set of the sequence that contradicts the choice from X, a 5-tuple of selves was created who are passionate enough to overrule u^* and guarantee that the c-choice from this set is the f-maximizer (while all other 5-tuples will be indifferent). Similarly, whenever along the sequence there is a set that contradicts the choice of the previous set, another 5-tuple of selves was created that overrules the preferences of all selves created in association with larger sets.

The above construction implies that if we permit the model to have n selves, any choice function (on any grand set of alternatives) having fewer than $\frac{n-1}{5}$ IIA-violations can be rationalized using this aggregator.

4.2 Main rationalizability result

The construction from the previous subsection can be generalized to any aggregator having the property that there exists $k \in Z_+$ such that there exists a triple-basis consisting of k selves that are arbitrarily close to being indifferent. As we showed above, passion-driven aggregators satisfy this requirement with $k = 5$. This property is relatively simple to check for a concrete aggregator, since it is defined for a three-element set. For scale-invariant aggregators, which satisfy the property that measuring utilities in a different unit does not change the ordering implied by the aggregator, checking the property is particularly simple, since it then suffices to construct one triple-basis which can be scaled as needed. For investigating how large is the set of aggregators satisfying the property, and for sufficient conditions for the property, see the next subsection.

Definition 4.1. *We say $\hat{U} \in \mathcal{U}(\{a, b, c\})$ is a triple-basis for f with respect to $\{a, b, c\}$ if $f(a, \{a, b\}, \{a, b, c\}, \hat{U}) > f(b, \{a, b\}, \{a, b, c\}, \hat{U})$, and $f(\cdot, A, \{a, b, c\}, \hat{U})$ is constant for all other $A \subseteq \{a, b, c\}$.*

Condition (Triple-solvability of f with k selves) There exists a triple $\{a, b, c\}$ and $k \in Z_+$ such that for every $\delta > 0$, there is a $U \in \mathcal{U}^k(\{a, b, c\})$ that is a δ-indifferent triple-basis for f with respect to $\{a, b, c\}$.

Triple-solvability with k selves implies that we can find a sequence of triple-bases containing k selves that converge to indifference.

Our main theorem applies to all aggregators satisfying P1-P6 and triple-solvability

Theorem 4.2. *Suppose f satisfies P1-P6 and is triple-solvable with k_f selves. Then, using n selves, f can rationalize any choice function c, defined on any finite grand set of alternatives X, that exhibits at most $\frac{n-1}{k_f}$ IIA-violations.*

The choice functions exhibiting no more than $\frac{n-1}{k_f}$ IIA-violations constitute a lower bound on behaviors that an aggregator f satisfying the conditions in the theorem can rationalize. The result can be restated such that if f satisfies P1-P6 and is triple-solvable with k_f selves then it can rationalize any choice function c with no more than $1 + k_f \cdot \text{IIA}(c)$ selves. That is, $1 + k_f \cdot \text{IIA}(c)$ is an upper bound on the number of selves (or complexity) required for rationalizing c with f. It is therefore evident that, in spite of having a structured form, any aggregator satisfying these properties can rationalize any choice function if sufficiently many selves are permitted by the model.

4.3 Sufficient conditions for the main result

A natural question is how large is the set of aggregators that satisfy the triple-solvability condition in Theorem 4.2. As for the aggregators featured in Section 2.3, which are closely related to models proposed in the existing literature, it is straightforward to show that all of them satisfy the condition. For example, permuting the alternatives a and b in the selves in the triple-basis for the passion-driven aggregator featured in the previous section works for passion-muting aggregators.[15]

Below we formally investigate how large the set of aggregators satisfying the triple-solvability condition is within the class of anonymous, additive and scale-invariant aggregators, and find uniform bounds for k_f. Formally, we impose the following additional structure.

P7 (Anonymity) For any $U \in \mathcal{U}^n(X)$ and any permutation $\pi : \left(R^X\right)^n \to \left(R^X\right)^n$ it holds that $f(\cdot,\cdot,X,U) = f(\cdot,\cdot,X,\pi(U))$.

P8 (Scale invariance) There is an invertible and odd $\phi : R \to R$ such that $f(\cdot,\cdot,X,\alpha U) = \phi(\alpha) f(\cdot,\cdot,X,U)\ \forall\ \alpha \in R$.

P9 (Additive separability) f is additively separable, i.e. $f(\cdot,A,X,U) = \sum_{i=1}^{n} g_i^A(u_i)$ $\forall\ X \in \mathcal{X}$, $A \subset X$ and $U \in \mathcal{U}^n(X)$.

Anonymity implies that the aggregation is symmetric with respect to selves. Scale-invariance implies that the ordering of different elements in the aggregation does not depend on the scale in which utilities are measured. Additive separability is a strengthening of P3, and is a common functional form assumption.[16]

Together, these properties imply that the aggregator $f(a,A,X,U)$ takes the form $\sum_{i=1}^{n} f(a,A,X,u_i)$, where $f(a,A,X,\alpha u) = \phi(\alpha) f(a,A,X,u)$.[17]

To measure the difference in aggregate utilities over any two alternatives within

[15]Contact the authors for triple-basis for aggregators featured in the other examples.

[16]To break this condition down, consider the following properties:
P9' (Strong Continuity) The ordering of elements over A implied by $f(\cdot,A,X,U)$ is continuous in U for every $A \subset X$.
P9" (Independence) For any $A \subset X$ and $a \in A$, and any $U, U', V', V'' \in \mathcal{U}$, $f(a,A,X,(U,V')) = f(a,A,X,(U',V'))$ implies $f(a,A,X,(U,V'')) = f(a,A,X,(U',V''))$.
Note that P9' implies P3, which requires the ordering of elements implied by f to be continuous in U only at indifferent selves. Then P1, P2 and P4, together with P9' and P9" imply P9, by *Debreu's aggregation theorem.* We do not provide the details here, referring the interested reader to Debreu (1959, Theorem 3) and also Maskin (1978).

[17]It actually suffices for our results that f is equivalent to such an aggregator (see the discussion on ordinal properties of our model in Subsection 2.1).

any choice set, we introduce the shorthand

$$f_{ac}(u) = f(a, \{a,c\}, \{a,b,c\}, u) - f(c, \{a,c\}, \{a,b,c\}, u),$$
$$f_a(u) = f(a, \{a,b,c\}, \{a,b,c\}, u)$$

for arbitrary $\{a,b,c\}$ and any single self $u \in \mathbb{R}^{\{a,b,c\}}$; as well as the shorthand $f_{ac}(U) = \sum_{i=1}^{n} f_{ac}(u_i)$ and $f_a(U) = \sum_{i=1}^{n} f_a(u_i)$ for a collection of selves U.

Suppose that there exists a positive constant γ and possibly menu-dependent constants δ_A such that the aggregator f takes the form $f(a, A, \{a,b,c\}, u) = \gamma u_i(a) + \delta_A$ for every a and every $A \subseteq \{a,b,c\}$. In this case, it is easy to see that although aggregate utility is affected cardinally by menu-dependence, it is not affected ordinally, and the resulting choice behavior is always rational. In particular, for any collection of selves U, knowing how the aggregator acts on any two pairs (for example, $f_{ab}(U)$ and $f_{bc}(U)$) one may immediately recover $f_{ab}(U)$ as the sum of these. To rule out such *degenerate* menu dependence, we introduce the following definition.

Definition 4.3. *If $U \in \mathcal{U}(\{a,b,c\})$, we say that $f \circ U$ is type-1 nondegenerate (on $\{a,b,c\}$) if $f_{ac}(U) \neq f_{ab}(U) + f_{bc}(U)$.*

As seen from the discussion above, nondegeneracy requires that f "stretch" utility differences in a manner that depends on the choice set. That is, preference intensity must be affected by the alternatives at hand.

Observe that for any $\{a,b,c\}$ and any $U \in \mathcal{U}(\{a,b,c\})$, $f \circ U$ generates a ranking of the alternatives in the triple: a is better than b in the triple if $f_a(U) \geq f_b(U)$. If $f \circ U$ rationalizes a choice function, there is a unique best element in the triple which is selected by the choice function. The *worst elements in the triple* are those $x \in \{a,b,c\}$ such that $f_x(U) \leq f_y(U)$ for all $y \in \{a,b,c\}$. Observe that $f \circ U$ also generates a ranking over alternatives from pairwise choice: a is better than b if $f_{ab}(U) \geq 0$. If the ranking generated from pairs is transitive, and $f \circ U$ generates a choice function, then there exists a unique *worst pairwise element.*

The next theorem establishes that if there exist selves U, defined on the triple $\{a,b,c\}$, for which $f \circ U$ is type-1 nondegerate and rationalizes some irrational behavior where the worst element according to pairwise choice "moves up" in the triple, then the aggregator is triple-solvable.

Theorem 4.4. Let f satisfy P1-P9. Consider any $\{a,b,c\} \in \mathcal{X}$, and suppose there exists $U \in \mathcal{U}(\{a,b,c\})$ such that $f \circ U$ is type-1 nondegenerate and rationalizes an irrational behavior in which the worst element over the pairs is not among the worst elements in the triple. Then there is k_f such that f is triple-solvable with k_f selves.

In particular, this implies that if $f \circ U$ is type-1 nondegenerate and satisfies one of the following two simple "extreme-switching" properties, then it is triple-solvable: (i) the worst element over the pairs is the best in the triple; or (ii) the best element over the pairs is the worst element in the triple. The first behavior corresponds to what GH call "third-place choice." The second occurs in one type of "second-place choice." If the aggregator can rationalize such preference shifts on a triple, it can rationalize any choice behavior on any grand set of alternatives; and with a fixed number of selves, the bounds of Theorem 4.2 apply.

The above result does not provide a bound on the number of selves in the triple-basis. The next theorem gives a uniform bound, under a different but related assumption on the aggregator using one self defined on the triple.

Definition 4.5. *If* $u \in \mathbb{R}^{\{a,b,c\}}$*, we say that* $f \circ u$ *is type-2 nondegenerate (on* $\{a,b,c\}$*) if* $[f_a(u) - f_b(u)] + [f_a(u) - f_c(u)] \neq f_{ab}(u) + f_{ac}(u)$.

This non-degeneracy condition rules out that the same linear relationship holds, for any possible self, between aggregated utilities given three-alternative and two-alternative sets.[18] These nondegeneracy conditions combined provide a uniform bound for the number of selves in a triple-basis.

Theorem 4.6. Suppose f satisfies P1-P9. Consider any $\{a,b,c\} \in \mathcal{X}$. If there exists a single self u on $\{a,b,c\}$ such that $f \circ u$ is type-1 and type-2 nondegenerate then f is triple-solvable with no more than 5 selves.

[18] It is easy to show, for example, that the passion-driven aggregator of Example 2.2 violates this condition of linear context-dependence whenever $g(\cdot)$ is nonlinear. Contact the authors for a proof.

5 Applications

5.1 Choice over menus: a generalized Strotzian model

In this section we apply our model and the results of Section 4 to examine the situation in which the DM is allowed to choose the menu from which she will pick an alternative. We will refer to the selection of a menu as the first stage of the decision problem, and posit the following testable restrictions on first-stage choice behavior. Denoting the grand set of alternatives by X, we assume that the DM has a preference relation $\succeq$ on $P(X) \times P(X)$ satisfying three simple axioms.

Axiom 1 (Preference Relation) $\succeq$ is complete and transitive

The preference $\succeq$ is a strict ordering on $\{a\}_{a \in X}$ if for all $a, b \in X$, $\{a\} \not\sim \{b\}$.

Axiom 2 (Strict Ordering) $\succeq$ is a strict ordering on $\{a\}_{a \in X}$

In the classical theory of choice, a set is assumed to be indifferent to its best element. Since then, various authors have relaxed this assumption by assuming, for example, that there are psychological costs to be borne by the introduction of unchosen but tempting elements, as in Gul and Pesendorfer (2001). Instead of such psychological costs, our model emphasizes inner conflict in choosing amongst alternatives. The set of alternatives may affect the chosen alternative in a manner that violates IIA. However, we retain the idea that the set is indifferent to the "best" element inside it, even if that element may not arise from a menu-independent ranking. That is, we posit the *Independence of Utility to Unchosen Alternatives* (IUUA): *taking as given whatever element is chosen*, the unchosen alternatives do not affect the well-being of the DM.

Axiom 3 (IUUA) For all $A \in P(X)$, there exists $a \in A$ such that $A \sim \{a\}$

Axiom 3 says that given a set of available menus and hence fixed first stage preferences of the DM, any decision maker maximizing these preferences is indifferent between the choice set A and getting just a, which we interpret as the element that is foreseen to be chosen from A.

This implies that for each prize $a \in X$, there is an equivalence class (let us call these classes "bins") and that each menu $A \in P(X)$ falls into one of these bins. Axioms 1-3 together ensure that we may uniquely define an *induced choice function* $c_{\succeq} : P(X) \to X$ by $c_{\succeq}(A) = a$ if $a \in A$ and $A \sim \{a\}$. We may then obtain the following representation theorem for choice over menus, which also provides a bound on the *second-stage subjective state space*.

Theorem 5.1. *$\succeq$ satisfies Axioms 1-3 if and only if there exist selves $U = (u_1, u_2, \dots, u_n) \in \mathcal{U}(X)$ and a utility function $W : X \to \mathbb{R}$ on prizes such that $\succeq$ is represented by the utility function $V : P(X) \to \mathbb{R}$ on sets, defined by*

$$V(A) = W\Big(\arg\max_{a \in A} f(a, A, X, U)\Big),$$

where f satisfies P1-P6 and triple-solvability with k selves, and $n \leq 1 + k \cdot \text{IIA}(c_{\succeq})$.

Proof. Because each menu is indifferent to the alternative chosen by the induced choice function, the DM's preferences over menus may be represented by a utility function $W(\cdot)$ over the alternatives in X. We may then use the result of Theorem 4.2 to rationalize the induced choice function. ∎

The representation may be interpreted as follows. When evaluating a choice set, the DM considers the various, possibly conflicting interests that will govern her choice from the set. These interests are represented by the selves U. The motivations that govern her choices from different menus need not be the same as W, which governs her choice over menus (although those motivations might be related to W). The DM simply picks the set from which the element foreseen to be chosen yields the greatest first stage utility. Consequently, the representation may be thought of as a generalized Strotzian preference (Strotz (1955)), where the DM chooses the best menu subject not to the choice of one self, but rather the choice maximizing the aggregate utility of multiple selves.[19]

The model implies that for any pair $\{a, b\}$, either $\{a, b\} \sim \{a\}$ or $\{a, b\} \sim \{b\}$. However, for larger sets, it may be that $A \cup B \succ A, B$ (interpreted as a preference for flexibility in Kreps (1979)), that $A, B \succ A \cup B$, or that $A \succ A \cup B \succ B$ (as in Gul and Pesendorfer (2001)'s betweenness, which they interpret in terms of costly self-control). The interpretation here is different: the DM is conflicted when she makes her choice from the menu, and depending on how she resolves the compromise among selves, might prefer a larger or smaller set that leads to a better choice according to the ex-ante utility W. How $A \cup B$ stands in relation to A and B provides information as to when the DM expects to be conflicted.

[19] We thank Eddie Dekel for suggesting this interpretation. We note that the above conception can reverse the logic in the branch of temptation and self-control literature begun by Gul and Pesendorfer (2001), but bears a relation to the separation of decision utility and experienced utility proposed by Kahneman, Wakker and Sarin (1997).

Observation 5.1. *If $A \cup B$ is not indifferent to either A or B then an IIA violation necessarily occurs in the induced choice function; and when an IIA violation occurs, the upper bound on the minimal number of states (in our interpretation, selves) required to rationalize the behavior increases.*

Using one of the *approximately triple-solvable* aggregators introduced in Section 6.1, each IIA violation in the induced choice function corresponds to one additional state. This is related to Dekel et al. (2001)'s result, where the subjective state space in a model of unforseen contingencies grows when there is additional desire for flexibility or self-control.[20] Here, "anticipated" IIA violations reveal additional conflicting motivations.

5.2 Microeconomic models of collective household choice

Empirical evidence on household demand strongly suggests that it cannot arise from the maximization of a single utility. An extensive literature examines the microeconomic implications of collective choice in households where each member is a utility maximizer; and in particular, a branch of this literature examines such models under the restriction of Pareto-efficient household behavior. One question addressed in this setting is, given a household demand function over N goods, when do there exist n utility functions $\{u_i\}_{i=1}^{n}$ and a continuously differentiable function μ of prices and income such that the demand arises from the weighted utilitarian maximization of $\sum_{i=1}^{n} \mu(\text{price,income}) u_i(\cdot)$ given the budget set (i.e., weights and preferences vary independently). Browning and Chiappori (1998) show that if there are N goods, then any demand data can be explained by an $(N-1)$-person household. In addition, to explain a given demand function using n people, it is necessary and sufficient that the rank of a certain matrix in a pseudo-Slutsky matrix decomposition be $n-1$, though without further restrictions there can be a continuum of explanatory n-person models (Chiappori and Ekeland (2006)).[21]

[20]It is also related to a trend seen in Gul and Pesendorfer (2005a) and Gul and Pesendorfer (2005b): the case of no self-control in Gul and Pesendorfer (2005a) ($A \cup B \sim A$ or $A \cup B \sim B$ for all A, B) can be rationalized with a single utility determining choice from the set, whereas the less restrictive Betweenness-based model of self-control in Gul and Pesendorfer (2005b), which only rules out violations in transitivity in an induced choice correspondence, can be rationalized with two utilities.

[21]The pseudo-Slutsky matrix is formally defined in Chiappori and Ekeland (2006); the rank condition they give, $SR(n-1)$, is that this matrix can be decomposed as the sum of a symmetric negative semi-definite matrix and another matrix of rank at most $n-1$. One intuition for the

To apply our framework in this context, we reinterpret selves as individuals of the household, and the aggregator as the mechanism that translates the individuals' preferences to household choice (this might be the outcome of a particular household bargaining procedure). Our approach differs in a number of ways from Browning and Chiappori (1998) and Chiappori and Ekeland (2006). First, the aggregator need not be weighted utilitarianism. Second, we address the question of rationalization by a concrete aggregator, while the above papers assume that the modeler does not know the underlying aggregation rule of the household, only that it belongs to the class of weighted utilitarian aggregators. Finally, we examine choice functions instead of demand functions. However, given that demand data is typically finite, suppose we denote by X the (finite) set of all available allocations, let each budget set correspond to a subset $A \subset X$, and identify the demand data with a function c that selects the allocation $c(A)$ in the budget set A. Then, rationalizing the demand data corresponds to rationalizing an *incomplete choice function*: c renders a choice to any subset A of X for some collection of subsets $\mathcal{A} \subset 2^X$, but data on choices from sets in $2^X/\mathcal{A}$ is missing. As we show below, our results can easily be extended to arbitrary incomplete choice functions.

Rationalizing an incomplete choice function c with aggregator f implies finding a set of selves U on X such that $f(c(A), A, X, U) > f(a, A, X, U)$ for all $a \in A/\{c(A)\}$ and $A \in \mathcal{A}$ (it does not matter what choices f and U imply from sets in $2^X/\mathcal{A}$). To see how our theorems generalize, observe that the only element of the construction that needs to be modified is the number of IIA violations: in this more general context we say that an IIA-violation is associated with choice set $A \in \mathcal{A}$ if there is a nested sequence of choice sets $A_1, A_2, ..., A_k$ such that $A_1 = X$, $|A_j| - |A_{j+1}| = 1 \ \forall\ j \in \{1, ..., k-1\}$, and $A_k = A$ for which the choice from A_k contradicts the choice from A_l for some $l < k$, and $A_{l'} \notin \mathcal{A}$ for any $l < l' < k$. It is easy to see that this definition reduces to the original one in case of no missing data. Once the definition of IIA(c) is modified accordingly, it can be shown that Theorem 4.2 holds (the proof is analogous).[22]

This means that for any aggregator satisfying our conditions, the demand data can be rationalized if there are sufficiently many people in the household. This

proof, which relies on exterior differential calculus, is that the Pareto-frontier for n people is $n-1$ dimensional, and weights and preferences can be varied independently.

[22]We note that IIA(c) for an incomplete choice function might be strictly less than IIA(c) for any completion $\widehat{c}$ of c. That is, it can be that any way of specifying choices for sets in $2^X/\mathcal{A}$ creates new IIA violations. Nevertheless, our theorems apply.

complements the result obtained in Browning and Chiappori (1998) and Chiappori and Ekeland (2006), in that even if the researcher knows how preferences in the household are aggregated, if the number of individuals in the (extended) household is large or unknown, then the model does not imply any testable restrictions on household demand. Our combinatorial approach also permits a simple lower bound on demand data that a household with a known number of individuals can generate, in terms of the number of IIA violations implied by the demand data.

6 Tightening the bounds on rationalizability

The bound on the set of rationalizable choice functions provided in our main results is not tight in general. Below we describe two methods of strengthening the results while keeping the basic features of our original construction, which in some cases lead to a tight bound. First, for some aggregators it is possible to find a tighter bound through a weakening of the triple-solvability requirement. Secondly, it may be possible to obtain a tighter bound by combining (or *collapsing*) some of the selves constructed, especially when the DM tends to make mistakes "in the same direction."

6.1 Approximate triple-solvability

For some aggregators a tighter upper bound can be provided for the minimum number of selves needed to rationalize a choice function, through a weakening of the triple-solvability requirement. In particular, it suffices for triple-solvability to hold only *approximately*, which can yield a triple-basis with a smaller number of selves. For ease of exposition we only state this property for additively separable aggregators.

Definition 6.1. *We say $\hat{U} \in \mathcal{U}(\{a,b,c\})$ is a (δ,ε)-approximate triple-basis for f with respect to $\{a,b,c\}$ if $f(a,\{a,b\},\{a,b,c\},\hat{U}) = f(b,\{a,b\},\{a,b,c\},\hat{U}) + \delta$ and $|f(x,A,\{a,b,c\},\hat{U}) - f(y,A,\{a,b,c\},\hat{U})| < \varepsilon$ for all other $A \subseteq \{a,b,c\}$ and $x,y \in A$.*

That is, a collection of selves U is a (δ,ε)-approximate triple basis for f if given choice set $\{a,b\}$ the aggregated utility of U for a is exactly δ higher than the aggregated utility of b, while U is ε-indifferent among all alternatives given every other choice set.

We say that an aggregator f is *approximately triple-solvable with k selves* if there is $\overline{\delta} > 0$ such that exists a (δ, ε)-approximate triple-basis with k selves for every $\delta < \overline{\delta}$ and $\varepsilon > 0$. That is, for approximate triple-solvability we do not require that the collection of selves in the triple is exactly indifferent between all elements in choice sets other than $\{a, b\}$, only that they can be arbitrarily close to being indifferent.

Theorem 4.2 can then be modified as follows.

Theorem 6.2. *Suppose f satisfies P1-P6 and P9, and is approximately triple-solvable with k_f selves. Then, for any finite set of alternatives X, and any choice function $c : P(X) \to X$ that exhibits at most $\frac{n-1}{k_f}$ IIA-violations, f can rationalize c with n selves.*

The proof is given in the Appendix.

To see why this result is powerful, take any aggregator of the form $f(a, A, X, U) = \sum_{i=1}^{n} h(\max_{a' \in A} u_i(a'))u_i(a)$, where $\lim_{x\to\infty} h(x)x = 0$. For example, consider a context-dependent version of relative utilitarianism,

$$f(a, A, X, U) = \sum_{i=1}^{n} \frac{u_i(a)}{1 + \big(\max_{b\in A} u_i(b)\big)^p}, \quad p > 1.$$

Under such an aggregator, the presence of an alternative with very high utility for a self means that self is given less say in the decision process (a "populist"-type model). This can be used to create a single-self approximate triple-basis u: let $u(a)$ and $u(b)$ such that $f(a, \{a, b\}, \{a, b, c\}, u) - f(b, \{a, b\}, \{a, b, c\}, u) = \delta$ (for small enough δ this is always possible), and let $u(c)$ be high enough so that u is ε-indifferent between any two elements given sets containing c. Theorem 6.2 then implies that the aggregator can rationalize all choice functions with no more than $n - 1$ IIA-violations, with n selves.

6.2 Collapsing triple-bases

Our construction allocates a different triple-basis (or approximate triple-basis) for every IIA-violation. However, there can be IIA-violations not contradicting each other, in which case parts of the associated triple-bases can be combined (or *collapsed*) together.

For example, recall the triple-basis we found for the passion-driven aggregator, and fix some alternative a. Observe that every time the choice of a from some set causes an IIA violation, the triple-basis constructed has a self u_5 in which a is preferred to $X \setminus \{a\}$, all elements of which are indifferent to each other. Under the passion-driven aggregator, all of the u_5 selves constructed when the choice was a can be collapsed into a single-self. That is, there would be a self for each distinct alternative whose choice causes an IIA violation, and four selves per violation in general. Consequently, "mistakes" in the same direction (e.g., always in the choice of a) can require fewer selves.

This effect is particularly pronounced when the triple-basis has only one self, as in the approximately triple-solvable aggregators introduced above. To illustrate this, consider the following example: let $x^* \in X$, and let $\succ_1$ and $\succ_2$ be strict orderings on X such that $x \succ_1 x^*$ and $x \succ_2 x^*$ for every $x \in X/\{x^*\}$, and $y \succ_1 x$ for $x, y \in X/\{x^*\}$ if and only if $x \succ_2 y$. Consider a decision-maker who from choice sets not containing x^* selects the best element according to $\succ_1$, but from choice sets containing x^* selects the best element according to $\succ_2$. This behavior describes, for example, a customer in a restaurant who chooses the tastiest item from a menu if the menu does not contain onion rings, while choosing the healthiest item in the presence of onion rings, because they are so greasy as to make the customer feel guilty about his eating habits.[23]

The above simple behavior generates a large number of IIA-violations if X is large.[24] However, these IIA-violations do not contradict each other: if choice from set B contradicts the choice from $A \supset B$, then there is no $B' \subset B$ such that the choice from B' contradicts the choice from B. As we show below, this can be used to merge all collections of selves into a single collection, drastically reducing the number of selves required to rationalize the above choice function.

Consider the context-dependent version of relative utilitarianism introduced in the previous subsection, which was shown to be approximately triple-solvable with a single self. Our construction calls for (i) creating a self whose utility function is in line with $\succ_2$; and (ii) creating a self for all sets associated with an IIA-violation,

[23]We thank Ran Spiegler for suggesting that we consider an example along these lines.

[24]The number of IIA-violations is $2^{n-1} - n - 1$: the choice from every set B having at least two elements and not containing x^* contradicts the choice from $B \cup \{x^*\}$.

such that the self attaches high enough utility to x^* such that the self becomes close enough to indifferent in the presence of x^*, and among the other alternatives allocates the highest utility to the choice from the given set. However, the latter selves can all be collapsed into a single self, such that the utility function of the self is in line with $\succ_1$ over $X/\{x^*\}$ (while keeping the utility of x^* at a level that makes the self nearly indifferent in the presence of x^*). This implies that the above choice function can be rationalized with two selves, which is obviously a tight bound.

7 Conclusion

The framework we propose in this paper provides a flexible environment for axiomatic investigation of multi-self models. As we pointed out, many of the models proposed in the existing literature can be translated into our framework such that the resulting aggregators satisfy the basic axioms we posited. However, there are other classes of aggregators that might be of interest, for example ordinal ones, which do not satisfy all our axioms. Our framework can still be useful to examine these aggregators, only some of our axioms need to be replaced by axioms that reflect the defining characteristics of the aggregators at hand. Furthermore, our set of axioms can also be supplemented with additional ones, leading to more specific classes of aggregators instead of the broad class of aggregation rules that we investigated in this paper, and hence to sharper predictions on implied choice with a fixed number of selves. We leave this direction, as well as extending our framework to dynamic settings, to future research.

Appendices

Appendix A: Proofs

Proof of Theorem 4.2. For an arbitrary choice function c we will construct a collection of $1 + k \cdot \mathrm{IIA}(c)$ selves which will be shown to rationalize c. This implies the claim in the theorem. In particular, we will construct k selves for each set with which an IIA-violation is associated, and an extra self for X.

Let $I_1 = \{A_1^1, ..., A_{i_1}^1\}$ be the subsets of X such that there is an IIA-violation associated with the set, but there is no proper subset of the set with which an IIA-violation is associated. For $j \geq 2$, let $I_j = \{A_1^1, ..., A_{i_{j+1}}^1\}$ be the subsets of X such that there is an IIA-violation associated with the set, but there is no proper subset of the set outside $\bigcup_{l=1}^{j-1} I_l$ with which an IIA-violation is associated. Let j^* be the largest j such that $I_j \neq \emptyset$.

We will now iteratively construct a k-tuple of selves for each set associated with an IIA-violation, starting with sets in I_1. Consider any k-tuple of selves $\bar{U}^1 = (\bar{u}_1^1, \ldots \bar{u}_k^1)$ that solves the triple $\{a, b, c\}$ (the existence of such a triple follows from triple-solvability). For every $A \subset I_1$, construct now the following collection of selves $U^A = (u_1^A, \ldots u_k^A)$:

$$u_i^A(x) = \begin{cases} \bar{u}_i^1(a) & \text{if } x = c(A) \\ \bar{u}_i^1(b) & \text{if } x \in A,\ x \neq c(A) \\ \bar{u}_i^1(c) & \text{if } x \notin A \end{cases}$$

for every $i = 1, ..., k$.

Suppose now that U^A is defined for every $A \in \bigcup_{k=1}^{j} I_k$ for some $j \geq 1$. Let U_k be the collection of selves $U_k = (U^{A_1^k}, ..., U^{A_{i_k}^k})$, for $k = 1, ..., j$. Let $\widehat{U}_j = (U_1, ..., U_j)$. By P4, there exists $\delta > 0$ such that for any k-tuple of δ-indifferent collection of selves U', $f(a, A, X, \widehat{U}_j) > f(b, A, X, \widehat{U}_j)$ implies $f(a, A, X, (\widehat{U}_j, U')) > f(b, A, X, (\widehat{U}_j, U'))$. Then by P3 and P6, we know $f(a, A, X, \widehat{U}_j, \widetilde{U}_1, ..., \widetilde{U}_m) > f(b, A, X, \widehat{U}_j, \widetilde{U}_1, ..., \widetilde{U}_m)$ implies the relation $f(a, A, X, (\widehat{U}_j, \widetilde{U}_1, ..., \widetilde{U}_m, U')) > f(b, A, X, (\widehat{U}_j, \widetilde{U}_1, ..., \widetilde{U}_m, U'))$ for any $\widetilde{U}_1, ..., \widetilde{U}_m$ collections of (exactly) indifferent selves.

Let now $I_{j+1} = \{A^1_1, ..., A^1_{i_{j+1}}\}$ be the subsets of X such that there is an IIA-violation associated with the set, but there is no proper subset of the set outside I_j with which an IIA-violation is associated. By triple-solvability with k selves, there is a δ-indifferent k-tuple of selves $\bar{U}^{j+1} = (\bar{u}^{j+1}_1, \dots \bar{u}^{j+1}_k)$ that solves the triple $\{a, b, c\}$. For every $A \subset I_{j+1}$, construct now the following collection of selves $U^A = (u^A_1, \dots u^A_k)$:

$$u^A_i(x) = \begin{cases} \bar{u}^{j+1}_i(a) & \text{if } x = c(A) \\ \bar{u}^{j+1}_i(b) & \text{if } x \in A,\ x \neq c(A) \\ \bar{u}^{j+1}_i(c) & \text{if } x \notin A \end{cases}$$

for every $i = 1, ..., k$. Let U_{j+1} be the collection of selves $(U_j, U^{A^1_1}, ..., U^{A^1_{i_{j+1}}})$.

The above procedure generates a collection of $k \cdot \text{IIA}(c)$ selves in j^* steps. Then by P3 and P4 there is $\delta_{j^*} > 0$ such that for any δ_{j^*}-indifferent u, $f(a, A, X, U_{j^*}) > f(b, A, X, U_{j^*})$ implies $f(a, A, X, (U_{j^*}, u)) > f(b, A, X, (U_{j^*}, u))$. Finally, construct one more self the following way: let $a_1 = c(X)$ and $a_k = c(X \setminus \{a_1, a_2, \dots a_{k-1}\})$ for $2 \leq k \leq n$. Construct $u^* : X \to \mathbb{R}$ such that $u^*(a_1) > u^*(a_2) > \cdots > u^*(a_n)$ and u^* is δ_{j^*}-indifferent.

We show the collection of selves $U_c \equiv (U_{j^*}, u^*)$ rationalize c with aggregator f.

Observation. First, note that for any set A with which there is an IIA violation associated, by the construction of U^A and by P1 and P5, $f(a, B, X, U^A) = f(b, B, X, U^A)\ \forall\ a, b \in B$ and B such that either $B/A \neq \emptyset$ or $c(A) \notin B$, and $f(c(A), B, X, U^A) > f(b, B, X, U^A) = f(b', B, X, U^A)\ \forall\ b, b' \in B/\{c(A)\}$ and B such that $B/A = \emptyset$ and $c(A) \in B$.

We will now show that the choice induced by f from any choice set is equal to the choice implied by c. First, note that this holds for X, since by the observation, $f(a, X, X, U^A) = f(b, X, X, U^A)$ for every $a, b \in X$ and every A with which there is an IIA-violation associated. Moreover, $f(c(X), X, X, u^*) > f(a, X, X, u^*)\ \forall\ a \in X/\{c(X)\}$ by P2. Then repeated application of P3 implies $f(c(X), X, X, U_c) > f(a, X, X, U_c)\ \forall\ a \in X/\{c(X)\}$.

Next, consider any $A \subsetneq X$ which causes an IIA violation. Suppose $A \in I_j$. The observation implies that for any $B \in (\bigcup_{l=1}^{j} I_l)/A$, $f(a, A, U^B) = f(a', A, U^B)\ \forall\ a, a' \in A$, and $f(c(A), A, X, U^A) > f(a, A, X, U^A)\ \forall\ a \in A$. Then repeated implication of P3 implies $f(c(A), A, X, U_j) > f(a, A, X, U_j)\ \forall\ a \in A$. By construction then

$f(c(A), A, X, U_c) > f(a, A, X, U_c)\ \forall\ a \in A$.

There are three cases to check for a set A that does not cause an IIA violation.

Case 1: For all $a \in A$, there is no $B \supset A$ such that $a = c(B)$. Then by construction $u^*(c(B)) > u^*(b)\ \forall\ b \in B/\{c(B)\}$. Moreover, by the observation, $f(b, B, X, U^A) = f(b, B, X, U^A)\ \forall\ b, b' \in B$ and A with which an IIA violation is associated. Repeated use of P3, together with P2, implies $f(c(B), B, X, U_c) > f(b, B, X, U_c)\ \forall\ b \in B$.

Case 2: There is a unique $a \in A$ such that for some $B \supset A$, $c(B) = a$. First we note that $a = c(A)$ is necessary, otherwise A would have caused an IIA violation. There are two subcases:

Case 2a: For every B such that $B \supset A$ and $c(B) = a$, B did not cause an IIA violation. This means that for all $B \supset A$, $c(B) \notin A \setminus \{c(A)\}$. So just like in Case 1, $u^*(c(B)) > u^*(b)\ \forall\ b \in B/\{c(B)\}$, and $f(b, B, X, U^A) = f(b, B, X, U^A)\ \forall\ b, b' \in B$ and A with which an IIA violation is associated. Hence, $f(c(B), B, X, U_c) > f(b, B, X, U_c)\ \forall\ b \in B$.

Case 2b: There is $B \supset A$ with $c(B) = a$ such that B caused an IIA violation. Consider any smallest such B, and suppose $B \in I_j$. By Observation 1, for any $A \in \bigcup_{l=1}^{j} I_l$ either $f(c(B), B, X, U^A) > f(b, B, X, U^A)\ \forall\ b \in B$, or $f(b, B, X, U^A) = f(b', B, X, U^A)\ \forall\ b, b' \in B$. But then repeated application of P3 implies that $f(c(B), B, X, U_j) > f(b, B, X, U_j)\ \forall\ b \in B$. By construction, $f(c(B), B, X, U_c) > f(b, B, X, U_c)\ \forall\ b \in B$.

Case 3: There exist at least two elements in A that have each been chosen in some superset. First, note that one of those elements must be $a = c(A)$, otherwise A would have caused an IIA violation. Let $\{b_i\}_i$ be the set of elements other than a such that $b_i \in A$ and $b_i = c(B_i)$ for some $B_i \supset A$. Drop any b_i's such that $B_i \subset B_m$ for some m and call the remaining set $\{b_j\}$. Because A did not cause an IIA violation by assumption, it must be that for each b_j there is A'_j such that $A \subset A'_j \subset B_j$ and $c(A'_j) \in A$. Because B_j does not contain any B_k, we know $c(A'_j) = a$. For each j there may be multiple such A'_j's; consider only the maximal A'_j with respect to the minimal B_j. Now by maximality, for any A'' such that $A'_j \subset A'' \subset B_j$, $c(A'') \notin A$. If there is A'' such that $c(A'') \in A'_j$, since $c(A'') \neq a$, by definition A'_j caused an IIA violation with respect to the first such A''. If for every A'' it is the case that $c(A'') \notin A'_j$, then once again A'_j caused an IIA violation with respect to B. Either way, since $c(A'_j) = a$, we added selves to ensure this choice for every j. This means that a should be the choice from A unless for some set B' between the smallest-sized A'_j and A we have $c(B') \in A \setminus \{a\}$ and selves were added. But by minimality of the B_j's there cannot be such a set. ∎

For compactness, we use the notation $x_1 = f_a(U) - f_b(U)$, $x_2 = f_b(U) - f_c(U)$, $x_3 = f_{ac}(U)$, $x_4 = f_{bc}(U)$, and $x_5 = f_{ab}(U)$. We prove the sufficiency conditions in the reverse order stated.

Proof of Theorem 4.6. We show a stronger result than stated: under type-1 non-degeneracy, if any one of the equations $2x_1+x_2-x_3-x_5=0$, $x_1+2x_2-x_3-x_4=0$, or $x_1-x_2+x_4-x_5=0$ fail then the aggregator is triple-solvable (with k_f at most $2+3|U|$).

The first column in the table lists the aggregate values for selves U. But by neutrality, we know that if we can generate the values in column 1, we can also generate the values in the 2nd column using the permutation $(bc)(a)$ over the alternatives, generate the values in the 3rd column using the permutation $(ab)(c)$ over the alternatives, and so on. By using duplication to evaluate each of the values $f \circ u$ and $f \circ u'$ each generated by a single self u and u', with the rankings given in the 6th and 7th headers, respectively, we can also generate the values in those respective columns.

$1:\ U$	$2:\ (bc)(a)$	$3:\ (ab)(c)$	$4:\ (abc)$	$5:\ (acb)$	$6: a \sim b \succ c$	$7: a \succ b \sim c$
x_1	x_1+x_2	$-x_1$	x_2	$-x_1-x_2$	0	x_1
x_2	$-x_2$	x_1+x_2	$-x_1-x_2$	x_1	x_1	0
x_3	x_5	x_4	$-x_5$	$-x_4$	x_1	x_1
x_4	$-x_4$	x_3	$-x_3$	x_5	x_1	0
x_5	x_3	$-x_5$	x_4	$-x_3$	0	x_1

Then, determinants of three possible 5×5 matrices, each composed of five of the columns above, may be calculated to obtain:

$$\mathrm{Det}(1|3|5|6|7) = x_1^2(x_1+2x_2-x_3-x_4)(2x_1+x_2-x_3-x_5)(x_3-x_4-x_5), \quad (1)$$
$$\mathrm{Det}(1|2|5|6|7) = x_1^2(2x_1+x_2-x_3-x_5)(x_3-x_4-x_5)(x_1-x_2+x_4-x_5), \quad (2)$$
$$\mathrm{Det}(2|3|4|6|7) = -x_1^2(x_1+2x_2-x_3-x_4)(x_3-x_4-x_5)(x_1-x_2+x_4-x_5). \quad (3)$$

To complete the proof, it suffices to show that there exists U such that defining $x_1, x_2, \ldots, x_5$ as above, one of the determinants in Equations (1)-(3) must be nonzero. If one of those determinants is nonzero, then we have find a vector $(c_1, c_2, c_3, c_4, c_5)$ such that the nonsingular matrix times $(c_1, c_2, c_3, c_4, c_5)$ is equal to $(0, 0, 0, 0, 1)$. Using scaling, each c_i can be pulled in so that the U corresponding to the i-th column is multiplied by c_i. The resulting set of selves provides a triple-basis (and therefore we can get triple solvability through scaling that triple-basis).

The proof is completed in light of the linear dependence of the equations $2x_1 + x_2 - x_3 - x_5 = 0$, $x_1 + 2x_2 - x_3 - x_4 = 0$, and $x_1 - x_2 + x_4 - x_5 = 0$: if any one of these fails, there must be a second which fails too. ∎

Proof of Theorem 4.4. By neutrality and symmetry of the type-1 nondegeneracy condition $x_3 - x_4 - x_5 \neq 0$, there are three types of choice behaviors we must examine to prove the result: one type of second-place choice (i), and both types of third place choice (ii-iii). The result then follows from the previous proof.
Cases 1: $a \succ_P b \succ_P c$ on the pairs, and $b \succ_T c \succeq_T a$ on the triple. If there is U such that $f \circ U$ rationalizes this behavior, then $x_3, x_4, x_5 > 0$ and $x_1 \leq 0$, $x_2 > 0$. Observe that $2x_1 + x_2 < 0$ since this is $f_a(U) - f_b(U) + f_a(U) - f_c(U)$. Therefore, $2x_1 + x_2 \neq x_3 + x_5$, as the RHS is positive.
Case 2: $a \succ_P b \succ_P c$ on the pairs, and $c \succ_T b \succeq_T a$ on the triple. That is, $x_3, x_4, x_5 > 0$, with $x_1 \leq 0$ and $x_2 < 0$. But as above, $2x_1 + x_2 \neq x_3 + x_5$, since the LHS is negative and the RHS is positive.
Case 3: $a \succ_P b \succ_P c$ on the pairs, and $c \succ_T a \succeq_T b$ on the triple. That is, $x_3, x_4, x_5 > 0$, with $x_1 \geq 0$, $x_2 < 0$. If we can find U such that $f \circ U$ rationalizes this behavior, then observe that $x_1 + 2x_2$ is negative since this is $f_a(U) - f_c(U) + f_b(U) - f_c(U)$. Hence $x_1 + 2x_2 \neq x_3 + x_4$ because the RHS is positive. ∎

Proof of Theorem 6.2. The only difference compared to the proof of Theorem 4.2 is in the construction of selves. Recall the definition of $(I_j)_{j=1,...,j^*}$ from the proof of Theorem 4.2. Let $\delta_1 \in (0, \overline{\delta})$. Define iteratively δ_j for $j \in \{2, ..., j^* + 1\}$ such that $\delta_j \in (0, \frac{\delta_{j-1}}{IIA(c)+1})$. Define a self u^X such that u^X is δ_{j^*+1}-indifferent and the preference ordering of the self is $c(X) \succ c(X/\{c(X)\}) \succ ...$ Let $\delta^{**} = \min_{x \neq y \in X,\ A \ni x,y} |f(x, A, X, u^X)| - |f(y, A, X, u^X)|$. Finally, let $\varepsilon \in (0, \frac{\delta^{**}}{|X|})$. Then for every $j \in \{1, ..., j^*\}$ and $A \in I_j$ construct a self u^A the following way: take a (δ_j, ε)-approximate triple-basis u, and let $u^A(c(A)) = u(a)$, $u^A(x) = u(b)\ \forall\ x \in A/\{c(A)\}$, and $u^A(x) = u(c)\ \forall\ x \in X/A$. Proving the collection of selves consisting of u^X and u^A for each $A \in \bigcup_{j=1}^{j^*} I_j$ rationalizes c is analogous to the proof in Theorem 4.2. ∎

Appendix B: Relaxing P6

Our main results can be extended to aggregators violating P6, that is, to aggregators that depend in a nontrivial way on alternatives unavailable in a given choice set. However, the appropriate definition of triple-solvability is more complicated.

The main complication arising in the absence of P6 is that triple-solvability needs to be defined on a general X, as opposed to just a triple $\{a, b, c\}$. It is convenient to introduce the following notation: for any triple $\{a, b, c\}$, any basic set of alternatives $X \supset \{a, b, c\}$, and any self u defined on $\{a, b, c\}$, define the set $E(u, X) = \{\widehat{u} : X \to \{u(a), u(b), u(c)\} | \widehat{u}(x) = u(x)\ \forall\ x \in \{a, b, c\}\}$. In words, $E(u, X)$ is the set of extensions of u from $\{a, b, c\}$ to X for which each element in $X/\{a, b, c\}$ receives the same utility as either a or b or c. Similarly, for any $U = (u_1, ..., u_m) \in \mathcal{U}(\{a, b, c\})$, let $E(U, X) = \{(\widehat{u}_1, ..., \widehat{u}_m) | \widehat{u}_i \in E(u_i, X)$ for all $i \in \{1, ..., m\}\}$.

Definition B.1. *We say $U \in \mathcal{U}(\{a, b, c\})$ is a universal triple-basis for f if for any $X \supset \{a, b, c\}$ the following holds: for all $\hat{U} \in E(U, X)$, $f(a, \{a, b\}, X, \hat{U}) > f(b, \{a, b\}, X, \hat{U})$, and $f(\cdot, A, X, \hat{U})$ is constant for all other $A \subseteq \{a, b, c\}$.*

A universal triple-basis solves the triple $\{a, b, c\}$ whenever the utilities of unattainable elements don't differ from utilities of elements in $\{a, b, c\}$, for all selves in the triple-basis. An aggregator f is *universally triple-solvable* if the following condition is satisfied.

Condition (Universal triple-solvability of f) There exists a triple $\{a, b, c\}$ and $k \in Z_+$ such that for every $\delta > 0$ there is a δ-indifferent $U \in \mathcal{U}^k(\{a, b, c\})$ constituting a universal triple-basis for f with respect to $\{a, b, c\}$.

It is easy to see that for aggregators satisfying P6, universal triple-solvability is equivalent to triple-solvability. If f satisfying P1-P5 is universally triple-solvable with k selves, then the same construction can be applied as in the proof of Theorem 4.2 to obtain an analogous lower bound on the set of choice functions that f can rationalize with a given number of selves. The proof of this result is analogous to the proof of Theorem 4.2 and hence omitted.

Theorem B.2. *Suppose f satisfies P1-P5 and is universally triple-solvable wrt to X with k_f selves. Then, using n selves, f can rationalize any choice function, on any grand set of alternatives X, that exhibits at most $\frac{n-1}{k_f}$ IIA-violations.*

References

Arrow, Kenneth J. and Hervé Raynaud, *Social Choice and Multicriterion Decision-Making*, Cambridge, Massachussetts: The MIT Press, 1986.

Benabou, Roland and Marek Pycia, "Dynamic Inconsistency and Self-Control: A PlannerDoer Interpretation," *Economics Letters*, 2002, *77*, 419424.

Bernheim, Douglas and Antonio Rangel, "Beyond Revealed Preference: Choice Theoretic Foundations for Behavioral Welfare Economics," *Working Paper*, 2007.

Browning, Martin and Pierre-André Chiappori, "Efficient Intra-Household Allocations: A General Characterization and Empirical Tests," *Econometrica*, 1998, *66*, 1241–1278.

Chatterjee, Kalyan and R. Vijay Krishna, "Menu Choice, Environmental Cues and Temptation: A 'Dual Self' Approach to Self-Control," *American Economic Journal: Microeconomics*, forthcoming.

Cherepanov, Victor, Tim Feddersen, and Alvaro Sandroni, "Rationalization," *Working Paper*, 2008.

Chiappori, Pierre-André and Ivar Ekeland, "The Microeconomics of Group Behavior: General Characterization," *Journal of Economic Theory*, 2006, *130*, 1–26.

Debreu, Gérard, "Topological Methods in Cardinal Utility Theory," *Cowles Foundation Paper 156*, 1959.

Dekel, Eddie, Barton Lipman, and Aldo Rustichini, "A Unique Subjective State Space for Unforeseen Contingencies," *Econometrica*, 2001, *69*, 891–934.

Dhillon, Amrita and Jean-Francois Mertens, "Relative Utilitarianism," *Econometrica*, 1999, *67*, 471–498.

Donahue, Eileen, Richard Robins, Brent Roberts, and Oliver John, "The Divided Self: Concurrent and Longitudinal Effects of Psychological Adjustment and Social Roles on Self-Concept Differentiation," *Journal of Personality and Social Psychology*, 1993, *64*, 834–846.

Evren, Ozgur and Efe Ok, "On the Multi-Utility Representation of Preference Relations," *Working Paper*, 2007.

Fudenberg, Drew and David Levine, "A Dual Self Model of Impulse Control," *American Economic Review*, 2006, *96*, 1449–1476.

Green, Jerry and Daniel Hojman, "Choice, Rationality, and Welfare Measurement," *Working Paper*, 2007.

Gul, Faruk and Wolfgang Pesendorfer, "Temptation and Self-Control," *Econometrica*, 2001, *69*, 14031436.

—— **and** —— , "The Revealed Preference Theory of Changing Tastes," *Review of Economic Studies*, 2005, *72*, 429–448.

—— **and** —— , "The Simple Theory of Temptation and Self Control," *Working Paper*, 2005.

Kahneman, Daniel, Peter Wakker, and Rakesh Sarin, "Back to Bentham? Explorations of Experienced Utility," *The Quarterly Journal of Economics*, 1997, *112*, 375–405.

Kalai, Gil, Ariel Rubinstein, and Ran Spiegler, "Rationalizing Choice Functions By Multiple Rationales," *Econometrica*, 2002, *70*, 24812488.

Kamenica, Emir, "Contextual Inference in Markets: On the Informational Content of Product Lines," *American Economic Review*, forthcoming.

Keeney, Ralph L. and Howard Raiffa, *Decisions with Multiple Objectives*, Cambridge, U.K.: Cambridge University Press, 1993.

Kivetz, Ran, Oded Netzer, and V. Srinivasan, "Alternative Models for Capturing the Compromise Effect," *Journal of Marketing Research*, 2004, *41*, 237–257.

Kochov, Asen, "The Epistemic Value of a Menu and Subjective States," *Working Paper*, 2007.

Kreps, David, "A Representation Theorem for 'Preference for Flexibility'," *Econometrica*, 1979.

Lachmann, Frank M., "How Many Selves Make a Person?," *Contemporary Psychoanalysis*, 1996, *32*, 595–614.

Lee, Leonard, On Amir, and Dan Ariely, "In Search of Homo Economicus: Preference Consistency, Emotion, and Cognition," *Working Paper*, 2007.

Manzini, Paola and Marco Mariotti, "Sequentially Rationalizable Choice," *American Economic Review*, 2007, *97*, 1824–1839.

Masatlioglu, Yusufcan and Efe Ok, "Rational Choice with Status Quo Bias," *Journal of Economic Theory*, 2005, pp. 1–29.

Maskin, Eric, "A Theorem on Utilitariansim," *The Review of Economic Studies*, 1978, *45*, 93–96.

May, Kenneth O., "Intransitivity, Utility, and the Aggregation of Preference Patterns," *Econometrica*, 1954, pp. 1–13.

Mitchell, Stephen, *Hope and Dread in Psychoanalysis*, New York: BasicBooks, 1993.

Power, M.J., "The Multistory Self: Why the Self is More Than the Sum of Its Autoparts," *Journal of Clinical Psychology*, 2007, *63*, 187–198.

Saari, Donald G., "Explaining All Three-Alternative Voting Outcomes," *Journal of Economic Theory*, 1999, *87*, 313–355.

Salant, Y. and A. Rubinstein, "Choice with Frames," *The Review of Economic Studies*, 2008, *75*, 1287.

Salant, Yuval, "Procedural Analysis of Choice Rules with Applications to Bounded Rationality," *Working Paper*, 2007.

Segal, Uzi, "Let's Agree That All Dictatorships Are Equally Bad," *Journal of Political Economy*, 2000, *108*, 569–589.

Sen, Amartya K., "Choice Functions and Revealed Preference," *The Review of Economic Studies*, 1971, *38*, 307–317.

____ , "Internal Consistency of Choice," *Econometrica*, 1993, *61*, 495–521.

Shafir, Eldar, Itamar Simonson, and Amos Tversky, "Reason-Based Choice," *Cognition*, 1993, *49*, 11–36.

Simonson, Itamar, "Choice Based on Reasons: The Case of Attraction and Compromise Effects," *Journal of Consumer Research*, 1989, *16*, 158–174.

____ **and Amos Tversky**, "Choice in Context: Tradeoff Contrast and Extremeness Aversion," *Journal of Marketing Research*, 1992, pp. 281–295.

Strotz, R.H., "Myopia and Inconsistency in Dynamic Utility Maximization," *Review of Economic Studies*, 1955, *23*, 165–180.

Tversky, Amos, "Intransitivity of Preferences," *Psychological Review*, 1969, *76*, 31–48.

____ **and Daniel Kahneman**, "Loss Aversion in Riskless Choice: A Reference-Dependent Model," *Quarterly Journal of Economics*, 1991, *106*, 1039–1061.

____ **and Itamar Simonson**, "Context-Dependent Preferences," *Management Science*, 1993, *39*, 1179–1189.