Falsifiability, complexity, and choice theories

Ádám Galambos*

May 2012

Abstract

During the past decade, an extensive literature on the testability of collective choice theories has developed. (Carvajal et al, 2004; Sprumont, 2000) We show that choice theories are falsifiable in a strong sense, as defined by the notion of *finite and irrevocable testability (FIT-ness)*, which was introduced by Herbert Simon and Guy Groen (Simon and Groen, 1973). Another approach to examining the empirical content of a theory is known as *Ramsey eliminability*. We use descriptive complexity theory to refine the notion of Ramsey eliminability for finite structures, and thereby introduce degrees of testability in a natural way. Our complexity-based notion of eliminability reveals that the theory of individual preference maximization is more testable than the theory of Nash equilibrium.

Keywords: Falsifiability; Descriptive complexity; Choice theory; Ramsey eliminability

1 Introduction

Karl Popper described the origins of his notion of *falsifiability* in his desire to "*distinguish between science and pseudo-science*." (Popper, 1962, p. 33) He also equated testability with falsifiability, adding, however, that there are degrees of testability. In economics, the notions of testability and falsifiability entered with Hutchison's methodological challenge (Hutchison, 1938) and revealed preference theory. In Paul Samuelson's words, "[i]n its narrow version the theory of "revealed preference" confines itself to a finite set of observable price-quantity competitive demand data, and

*Department of Economics, Lawrence University, Appleton, Wisconsin. Email: adam.galambos@lawrence.edu

attempts to discover the full empirical implications of the hypothesis that the individual is consistent." (Samuelson, 1953) This consistency is embodied in some version of the Strong Axiom of Revealed Preference. Revealed preference theory serves as a methodological foundation for consumer theory, whose empirical content is demarcated by revealed preference conditions. During the past decade, an extensive literature on the testability of collective choice theories has developed. (Carvajal et al, 2004; Sprumont, 2000) The contribution of this paper is examine the falsifiability and Ramsey eliminability of certain choice theories, and to propose a refinement of the notion of Ramsey eliminability that is shown to be useful in the context of choice theories.

Since Popper, philosophers of science have turned to model theory to formalize the notions of theory and falsifiability. A theory is described as a set of sentences in a formal language $\mathcal{T}$, which particular structures (models) stated in the language $\mathcal{T}$ may or may not satisfy.[1] In order to formalize the notion of observation, the language has two parts: a more restrictive language of observation $\mathcal{O}$, and the full language $\mathcal{T}$ that also includes theoretical terms in addition to those used to describe observations. Several criteria by which theories may be judged have been proposed. The simplest of these calls a theory *falsifiable* if there exists an $\mathcal{O}$-structure that cannot be extended into a $\mathcal{T}$-structure in such a way that it satisfies the theory. This is, indeed, a very basic and very weak requirement, so a more stringent version has been proposed. *Strong falsifiability* of a theory means that *every* $\mathcal{O}$-structure that cannot be extended into a $\mathcal{T}$-structure to satisfy the theory has an $\mathcal{O}$-substructure that cannot be embedded isomorphically into an $\mathcal{O}$-structure that is extendable into a $\mathcal{T}$-structure to satisfy the theory. Strong satisfiability means that every set of observations that is inconsistent with the theory contains a subset of observations that could not possibly be "saved" by any additional observations. Thus every set of observations inconsistent with the theory contains irrevocable evidence of that inconsistency. In fact, as we will see below, strong falsifiability is equivalent to *irrevocable testability.*

Another approach to examining the empirical content of a theory is to study the *Ramsey eliminability* of its theoretical terms. (Benthem, 1978) This entails determining whether the theory can be stated entirely in the language of observation, eliminating all theoretical terms. This enterprise resembles very closely the revealed preference approach in economics, where revealed preference theory attempts to reframe consumer theory in terms of the language of choice alone, without making use of the language of preference or utility.

Herbert Simon and Guy Groen (Simon and Groen, 1973) introduced the notion of

[1]See below for formal definitions. Structures in the language $\mathcal{L}$ are called $\mathcal{L}$-structures for short.

finitely and irrevocably testable (FIT) theories. FIT-ness is a strengthening of strong falsifiability in that they require that a $\mathcal{T}$-structure that does not satisfy the theory have a *finite* $\mathcal{O}$-substructure that cannot be extended into a $\mathcal{T}$-structure satisfying the theory. Moreover, such a falsifying $\mathcal{O}$-substructure must be irrevocable in the sense that it cannot be extended into a larger $\mathcal{O}$-structure that *can* be extended into a $\mathcal{T}$-structure satisfying the theory.

The purpose of this paper is two-fold: to examine the falsifiability of choice theories using the notion of *FIT*-ness, and to propose a complexity-based perspective on Ramsey eliminability. We introduce a (weaker) notion of a substructure appropriate to the structure of observations in choice theory, and show that choice theories are *FIT*. We refine the notion of Ramsey eliminability for finite structures, and thereby introduce degrees of testability in a natural way. Using the theories of individual preference maximization and Nash equilibrium, we show the usefulness of our new notion. The complexity-based view is appropriate only when the restriction to finite structures is acceptable, but in such cases, as in choice theory, it allows one to make important distinctions.

We do not argue that consumer theory or any other choice theory would have to be discarded on the grounds that it does not satisfy one or another falsifiability criterion. In fact, several philosophers of economic science have argued that a fasificationalist view is not appropriate for economics, for example because of the unreliability of data (see Hausman, 1992). At the same time, with the significant development of experimental economics, empirically testing choice theories is becoming a more realistic enterprise. In addition, a lack of falsifiability would surely not be insignificant, and may indicate that implications derived from economic choice theories rely on theoretical constructs that go well beyond what observable phenomena could possibly imply.[2] Moreover, our refinement of the notion of Ramsey eliminability gives meaning to Popper's statement that testability can be a matter of degree. This allows for a more sophisticated approach to testability than the binary notion of (strong) falsifiability or *FIT*-ness.

Our approach in this paper is syntactic rather than semantic. Though Rynasiewicz (1983) argues that a semantic approach is more fruitful generally, in the case of choice theory the syntactic approach works. More importantly, one of the contributions of this paper is the refinement of the notion of eliminability based on descriptive (and computational) complexity considerations. Thus the syntactic form of theories is at the center of our analysis.

Criteria of falsifiability have been applied to consumer theory in previous works. Boland (1981) argues that consumer theory is not falsifiable because it involves an

[2]For a variety of perspectives on the methodology of economics, see Backhouse (1994).

existential statement ("there exist binary relations such that ..."). Mongin (1986) points out that revealed preference theory proves Boland's contention wrong, and that consumer theory can be reformulated as a universal statement. Moreover, he points to the possibility that the more important distinction may be between the expressibility of a statement in second-order or first-order logic. We add refinement and further support to this view through our general proposal regarding a complexity-based view of testability, and show the usefulness of this perspective through the example of individual preference maximization and Nash equilibrium.

2 Falsifiability and *FIT*-ness

2.1 Definitions

The following definitions are standard—see, for example, Immerman (1999). We do not spell out standard notions in detail, such as the meaning of "truth," but we follow standard usage that the reader may find in any introduction to logic.

Definition 1. A *relational language* is a tuple

$$\mathcal{L} = \langle R_1^{a_1}, \ldots, R_r^{a_r}, c_1, \ldots, c_s \rangle \tag{1}$$

together with second-order predicate logic, where $R_i^{a_i}$ is a relation of arity a_i, and the c_i are constant symbols.

We use the implication "$\Rightarrow$" and bi-implication " $\iff$ " symbols in the conventional sense. The second-order quantifiers will be denoted by "$\forall_2$" and "$\exists_2$."

Definition 2. An $\mathcal{L}$-*structure* is a tuple

$$L_{\mathcal{L}} = \langle U, R_1^L, \ldots, R_r^L, c_1^L, \ldots, c_s^L \rangle, \tag{2}$$

where the finite nonempty set U is the *universe*, and each R_i^L is a relation of arity a_i on U. For each constant symbol c_i of $\mathcal{L}$, there is a constant $c_i^L \in U$.[3] Sometimes we will refer to R_i^L as the $L_{\mathcal{L}}$-interpretation of $R_i^{a_i}$, and we extend this usage to sentences in $\mathcal{L}$ as well. We will occasionally omit the language-subscript and refer to a structure L, whenever the language used is clear from the context.

[3]Sometimes an $\mathcal{L}$-structure is called a model for $\mathcal{L}$.

The language of observation for choice theory for agents $N = \{1, \ldots, n\}$ is

$$\mathcal{O}^n = \langle S_1^1, \ldots, S_n^1, E^{n+1}, A^1, B^2, C^2 \rangle, \tag{3}$$

where each S_i^1 is a "universal strategy set" for i, the relation E^{n+1} is an outcome function[4] assigning an element of the universe to every strategy profile in $S_1^1 \times \cdots \times S_n^1$ (see example 2 below), the unary relation A^1 is the index set of observations, the binary relation B^2 describes choice sets ("budgets"), and the binary relation C^2 describes choices for each choice set. The language does not have any constant symbols. The language of theory is

$$\mathcal{T}^n = \langle S_1^1, \ldots, S_n^1, E^{n+1}, A^1, B^2, C^2, R_1^2, \ldots, R_n^2 \rangle, \tag{4}$$

where the R_i^2 are binary relations representing preferences. For these relations, we will use the notation xR_i^2y instead of $(x, y) \in R_i^2$.

Example 1. Consider the following individual choice data. A decision maker chooses from three-member subsets of the set $\{a, b, c, d\}$, and her choice is always the letter that comes first in the alphabet of those available. An $\mathcal{O}^n$-structure $\hat{O}$ describing this situation would have: the universe

$$\hat{U} = \{1, 2, 3, 4, 5, 6, 7, 8, a, b, c, d\};$$

the budget index set

$$A^{\hat{O}} = \{1, 2, 3, 4\},$$

the strategy set

$$S_1^{\hat{O}} = \{5, 6, 7, 8\},$$

with the remaining four elements of the universe ($\{a, b, c, d\}$) corresponding to choice alternatives via the outcome function

$$E^{\hat{O}} = \{(5, a), (6, b), (7, c), (8, d)\},$$

the budget relation

$$B^{\hat{O}} = \{(1, a), (1, b), (1, c), (2, a), (2, b), (2, d), (3, a), (3, c), (3, d), (4, b), (4, c), (4, d)\}$$

identifying the four budgets; and the choice relation

$$C^{\hat{O}} = \{(1, a), (2, a), (3, a), (4, b)\}.$$

[4] Even though we model the outcome function as a relation, we will refer to it as the outcome function for consistency with standard game theoretic usage.

Example 2. We observe two players in a normal form game. One has strategy set $\{a, b\}$, the other has strategy set $\{A, B\}$. We observe them choose (a, A) from the full 2×2 game, and (b, B) from the game where the first player is restricted to choosing b. An $\mathcal{O}^n$-structure $\hat{O}$ describing this situation would have: the universe

$$\hat{U} = \{1, 2, a, b, A, B, (a, A), (a, B), (b, A), (b, B)\},$$

the "budget"[5] index set

$$A^{\hat{O}} = \{1, 2\},$$

the strategy sets

$$S_1^{\hat{O}} = \{a, b\},\ S_2^{\hat{O}} = \{A, B\},$$

the outcome function

$$E^{\hat{O}} = \{(a, A, (a, A)), (a, B, (a, B)), (b, A, (b, A)), (b, B, (b, B))\},$$

the "budget" relation

$$B^{\hat{O}} = \{(1, (a, A)), (1, (a, B)), (1, (b, A)), (1, (b, B)), (2, (b, A)), (2, (b, B))\}$$

identifying the two game forms; and the choice relation

$$C^{\hat{O}} = \{(1, (a, A)), (2, (b, B))\}.$$

Definition 3. An $\mathcal{L}$*-theory* is a set of sentences in the language $\mathcal{L}$. An $\mathcal{L}$-structure M satisfies the $\mathcal{L}$-theory Σ if the M-interpretation of every sentence in Σ is true in M.

Assumption 1. Every theory we consider will include the logical axioms. In addition, every theory we consider will include the following sentence χ, which states the axioms defining a choice-theoretic structure:

$$\chi := \Big[\forall a, y\big[(a, y) \in C^2 \ \Rightarrow\ [a \in A^1 \ \wedge\ y \notin \bigcup_{i \in \{1,\ldots,n\}} S_i^1 \cup A^1]\big]\Big] \wedge \tag{5a}$$

$$\Big[\forall a, y\big[(a, y) \in B^2 \ \Rightarrow\ [a \in A^1 \ \wedge\ y \notin \bigcup_{i \in \{1,\ldots,n\}} S_i^1 \cup A^1]\big]\Big] \wedge \tag{5b}$$

$$\Big[\forall x\ \forall i_{i \in \{1,\ldots,n\}}[x \in S_i^1] \ \Rightarrow\ [x \notin A^1]\Big] \wedge \tag{5c}$$

$$\Big[\forall x_1, \ldots, x_{n+1}\big[(x_1, \ldots, x_{n+1}) \in E^{n+1}\big] \ \Rightarrow\ \big[[\forall i_{i \in \{1,\ldots,n\}} x_i \in S_i^1] \wedge [x_n \notin \bigcup_{i \in \{1,\ldots,n\}} S_i^1 \cup A^1]\big]\Big]. \tag{5d}$$

[5] We continue to refer to the alternatives available in an observation as "budgets," even though, in the case of multiple players, these are game forms. The analogy with budgets as available alternatives is, we hope, transparent.

(5a) says that the binary relation C^2 describing choice has an element of the index set for observations as its first component, and as its second component it has an element of the universe that is not from the index set for observations or the universal strategy sets. (5b) says the same thing for the binary relation defining budgets. (5c) says that the strategy sets are disjoint from the index set for observations, and (5d) says that the first n arguments of the relation defining the outcome function are from the appropriate universal strategy sets, and the last argument is not from the index set for observations or the universal strategy sets, i.e., it is from the set of alternatives.

Remark 1. Our language $\mathcal{O}^n$ of choice theory is similar to the *language of group choice* $\mathcal{F}$ introduced by Chambers et al (2010). However, there are several differences, one of which is significant. With the language $\mathcal{F}$ defined by Chambers et al (2010), an $\mathcal{F}$-structure, which is their description of the set of observations, lists every subset of every agent's strategy set. This means that the description of an $\mathcal{F}$-structure grows exponentially in the number of strategies, regardless of the number of observations. Because we will study questions of computational complexity as well, such a description would not be acceptable. Our description, using $\mathcal{O}^n$-structures, does not suffer from this problem.

Remark 2. In this paper, we are interested only in choice theoretic structures, so we will only consider $\mathcal{O}^n$- and $\mathcal{T}^n$-structures that satisfy χ above in (5). Because we assume throughout that outcome functions are observable, every falsifiability notion we consider will be *relative to choice structures satisfying* χ. (See Definition 7 below.) This simply means that we do not address the question of testing whether a particular structure is a choice theoretic structure or not. Rather, we are interested in whether a choice theory is falsifiable or not.

In our application, an $\mathcal{O}^n$-structure will describe a set of choices, and a $\mathcal{T}^n$-structure will describe a set of choices together with the theory of preferences explaining those choices.

Definition 4. A $\mathcal{T}^n$-structure T can be *restricted* to an $\mathcal{O}^n$-structure by simply discarding the preference relations R_i^T. Conversely, an $\mathcal{O}^n$-structure O can be *extended* to a $\mathcal{T}^n$-structure by simply adding binary relations on the universe of O defining preferences. In these cases we speak of a *restriction* of a $\mathcal{T}^n$-structure in $\mathcal{O}^n$, and an *extension* of an $\mathcal{O}^n$-structure in $\mathcal{T}^n$.

To consider falsifiability, we need to define the notion of an $\mathcal{O}^n$-substructure. Intuitively, a substructure describes a subset of a set of observations. In the case of choice theory, the set defined by the unary relation A^1 serves as an index set for observations, so it is natural to define a subset of observations by a subset of A^1.

Definition 5. A *substructure* of an $\mathcal{O}^n$-structure $O = \langle U, S_1^O, \ldots, S_n^O, E^O, A^O, B^O, C^O \rangle$ is an $\mathcal{O}^n$-structure $O' = \langle U \backslash (A^O \backslash A^{O'}), S_1^O, \ldots, S_n^O, E^O, A^{O'}, B^{O'}, C^{O'} \rangle$ where $A^{O'} \subseteq A^O$ and the relations $B^{O'}$ and $C^{O'}$ are the restrictions of B^O and C^O to $A^{O'}$.[6]

Definition 6. Let Σ be a $\mathcal{T}^n$-theory. An $\mathcal{O}^n$-structure O is *extendable to satisfy* Σ if it can be extended to a $\mathcal{T}^n$-structure that satisfies Σ.

Definition 7. Let Σ be a $\mathcal{T}^n$-theory including the axiom χ. Σ is *falsifiable relative to* χ if there exists an $\mathcal{O}^n$-structure satisfying χ that is not extendable to satisfy Σ. We say that Σ is *strongly (finitely) falsifiable relative to* χ if every $\mathcal{O}^n$-structure satisfying χ that is not extendable to satisfy Σ has a (finite) substructure that cannot be isomorphically embedded into an $\mathcal{O}^n$-structure that satisfies χ and is extendable to satisfy Σ.

Definition 8. Let Σ be a $\mathcal{T}^n$-theory including the axiom χ. Σ is *irrevocably testable relative to* χ if the property of extendability to satisfy Σ is preserved under $\mathcal{O}^n$-substructures satisfying χ. That is, there does not exist an $\mathcal{O}^n$-structure satisfying χ that is extendable to satisfy Σ, and has a substructure satisfying χ that is not extendable to satisfy Σ.

As Rynasiewicz (1983) remarks, strong falsifiability is equivalent to irrevocable testability. A theory that is strongly finitely falsifiable relative to χ will simply be called *FIT relative to* χ. The notion of *FIT*-ness (finite and irrevocable testability) was introduced by Simon and Groen (1973). Our relativization of the notion of *FIT*-ness follows the notion of relative falsifiability introduced by Rynasiewicz (1983).

2.2 *FIT*-ness of choice theories

We present a general formulation of choice theory in the language $\mathcal{T}^n$.

$$\begin{aligned}
\Psi^n(\varphi) := & \big[\forall a, y \, \big[[(a,y) \in C^2] \Rightarrow [\forall z_{(a,z)\in B^2} \, \varphi]\big]\,\big] \wedge && \text{(6a)}\\
& \big[\forall a, y \, \big[[(a,y) \in B^2] \wedge [(a,y) \notin C^2] \Rightarrow [\exists z_{(a,z)\in B^2} \, \neg\varphi]\big]\,\big] \wedge && \text{(6b)}\\
& \big[\forall i_{i\in\{1,\ldots,n\}} \forall x, y_{x,y\notin A^1} [x R_i^2 y] \vee [y R_i^2 x]\big] \wedge && \text{(6c)}\\
& \big[\forall i_{i\in\{1,\ldots,n\}} \forall x, y, z_{x,y,z\notin A^1} [x R_i^2 y \wedge y R_i^2 z] \Rightarrow x R_i^2 z\big]. && \text{(6d)}
\end{aligned}$$

[6]Note that the relations S_i^O and E^O are defined on $U \backslash A^O$ by (5c) and (5d), and so need not be restricted.

The first-order sentence $\varphi(y, z, S_1^1, \ldots, S_n^1, E^{n+1}, R_1^2, \ldots, R_n^2)$ defines the solution concept, and contains occurrences of the universal strategy sets S_i^1, the outcome function E^{n+1}, the preference relations R_i^2, unbound occurrences of y and z, as well as bound occurrences of other variables. Note that φ is not allowed to depend on A^1, so the solution for a particular budget cannot depend on other budgets that may or may not be part of the choice structure. (6a) says that for every observation, elements that are observed chosen are solutions, as defined by φ. (6b) says that for every observation, elements that are not observed chosen are not solutions, as defined by φ. (6c) says that each agent's preference relation is total, and (6d) that each preference relation is transitive. Thus the theory $\Psi^n(\varphi)$ defines φ-rationality for n decision makers. Note that $\Psi^n(\varphi)$ is, in fact, a first-order sentence.

To define *strict* φ-rationality, simply replace (6c) by

$$\big[\forall i_{i\in\{1,\ldots,n\}} \forall x, y_{x,y\notin A^1, x\neq y}[xR_i^2 y] \vee [yR_i^2 x]\big] \tag{6c'}$$

and add another conjunct expressing asymmetry and non-reflexivity of preferences:

$$\big[\forall i_{i\in\{1,\ldots,n\}} \forall x, y_{x,y\notin A^1} \neg\big[[xR_i^2 y] \wedge [yR_i^2 x]\big]\big]. \tag{6e}$$

The resulting theory (the conjunction of (6a), (6b), (6c'), (6d), and (6e)) will be denoted by $\Psi^n(\text{str-}\varphi)$.

In some applications, *partial* rationalizability is of interest. This theory can be obtained from $\Psi^n(\varphi)$ simply by omitting (6b), and will be denoted by $\Psi^n_{sub}(\varphi)$. In this case, what is observed as chosen is understood to be a subset of the φ-solutions, not the entire set of φ-solutions. (This notion has been called *sub-semirationality* by Matzkin and Richter (1991).) Symmetrically, in some applications what is observed as chosen is understood to be a superset of φ-solutions. The theory for this notion (called *supra-semirationality* by Matzkin and Richter (1991)) is obtained from $\Psi^n(\varphi)$ simply by omitting (6a), and will be denoted by $\Psi^n_{sup}(\varphi)$. The strict versions of these theories, $\Psi^n_{sub}(\text{str-}\varphi)$ and $\Psi^n_{sup}(\text{str-}\varphi)$, are defined analogously.

Example 3. To obtain the theory of individual preference maximizing choice, let $n = 1$ and let φ be "$yR_1^2 z$." In the case of individual preference maximizing choice, there is no role for the strategy set, and so we may define rationality without reference to S_1^1.

Example 4. To obtain the theory of Nash equilibrium, let φ be the sentence φ_{Nash} defined by

$$\forall i_{i\in\{1,\ldots,n\}} \Big[\Big[\forall j_{j\neq i} \exists x_{j_{x_j\in S_j^1}} \exists x_i, x'_{i\; x_i, x'_i\in S_i^1} \\ \big[(x_1, \ldots, x_n, y) \in E^{n+1} \;\wedge\; (x_1, \ldots, x'_i, \ldots, x_n, z) \in E^{n+1}\big]\Big] \;\Rightarrow\; yR_i^2 z\Big].$$

Example 5. Let $\mathcal{C}$ denote a set of subsets of $\{1,\ldots,n\}$. In a $\mathcal{C}$-strong Nash equilibrium, no coalition in $\mathcal{C}$ can jointly deviate to make each member of the coalition better off. To obtain the theory of $\mathcal{C}$-strong Nash equilibrium, let φ be

$$\forall S_{S\in\mathcal{C}}\Big[\Big[\forall j_{j\notin S}\exists x_{j\,x_j\in S_j^1}\forall i_{i\in S}\exists x_i,x'_{i\;x_i,x'_i\in S_i^1} \\ \big[(x_1,\ldots,x_n,y)\in E^{n+1}\;\wedge\;((x_j)_{j\notin S},(x'_i)_{i\in S},z)\in E^{n+1}\big]\Big]\;\Rightarrow\;\exists i_{i\in S}\neg[zR_i^2y]\Big].$$

Note that $\mathcal{C}$ is not part of the language $\mathcal{T}^n$, and it is not necessary to describe it, or to quantify over subsets of it. In any particular case of $\mathcal{C}$, the sentence φ involves conjunctions and disjunctions that could be written without reference to $\mathcal{C}$ or its subsets. The quantifiers used on $\mathcal{C}$ or its members in the sentence above are shorthand for conjunctions or disjunctions, used only for the sake of better comprehensibility. Because we state all theories for a fixed number n of individuals, in the sentence above quantification over $\mathcal{C}$ or S could be substituted with conjunctions or disjunctions.

Theorem 1. *For a first order solution φ, all choice theories $\Psi^n(\varphi),\Psi^n_{sub}(\varphi),\Psi^n_{sup}(\varphi)$ and their strict versions are FIT relative to χ. In particular, the theories of individual preference maximizing, Nash equilibrium, and $\mathcal{C}$-strong Nash equilibrium are FIT relative to χ.*

Proof. In all that follows, we assume that every structure is a choice structure, i.e., that it satisfies χ as stated in (5). First we show irrevocability. Intuitively, this is clear: if there exist preferences that rationalize a set of observations, then these same preferences also rationalize any subset of those observations. More precisely, if an $\mathcal{O}^n$-structure

$$\hat{O}=\langle\hat{U},S_1^{\hat{O}},\ldots,S_n^{\hat{O}},E^{\hat{O}},A^{\hat{O}},B^{\hat{O}},C^{\hat{O}}\rangle$$

can be extended to satisfy $\Psi^n(\varphi)$, then, by definition, there exist preferences $\hat{R}_1,\ldots,\hat{R}_n$ such that the $\mathcal{T}^n$-structure

$$\hat{T}=\langle\hat{U},S_1^{\hat{O}},\ldots,S_n^{\hat{O}},E^{\hat{O}},A^{\hat{O}},B^{\hat{O}},C^{\hat{O}},\hat{R}_1,\ldots,\hat{R}_n\rangle$$

satisfies $\Psi^n(\varphi)$. Consider a substructure O' of O, and extend it to a $\mathcal{T}^n$-structure $\hat{T}'$ with the preference relations $\hat{R}_1,\ldots,\hat{R}_n$. It is immediate that (6c) and (6d) will be satisfied by $\hat{T}'$. (6a) and (6b) will also hold for $\hat{T}'$ because the budget and choice relations are unchanged for all $x\in A^{O'}$, and so these axioms will continue to hold for all $x\in A^{O'}$.

As (Rynasiewicz, 1983, p. 230) notes, any first-order theory that is irrevocably testable is also finitely irrevocably testable by the Completeness Theorem for first-order logic. Thus $\Psi^n(\varphi)$ is *FIT*.

The arguments for $\Psi^n_{sub}(\varphi),\Psi^n_{sup}(\varphi)$ and their strict versions are analogous. □

Remark 3. An even stronger notion of falsifiability of a theory Σ could require that any $\mathcal{T}^n$-structure that does not satisfy Σ have an $\mathcal{O}^n$-substructure that is inconsistent with Σ (see (Rynasiewicz, 1983, footnote 1)). This requirement is generally too strong, as Rynasiewicz (1983) notes, and would not be satisfied by choice theories. Indeed, any structure that has no observations (that is, an empty A^1 relation) or very incomplete observations, could have intransitive preferences on alternatives that are not part of any observation, and this intransitivity could never be detected.

Remark 4. Our notion of a substructure is more restrictive than the standard notion that takes any subset of the universe and restricts the relations of the structure to that subset. Thus our result that choice theory is irrevocably testable is logically weaker than the statement that "choice theory is closed under *substructures in the standard sense*." We believe, however, that in the case of choice theory, our notion of a substructure is the natural one. An observation, in this context, is a set of choice possibilities for the decision makers, together with choice(s) made by them. Since a substructure is meant to describe a subset of observations, our definition of a substructure as a subset of the observations follows.

Remark 5. Our Theorem 1 establishes that choice theories are falsifiable in a strong sense. Our result is logically neither stronger nor weaker than those in Chambers et al (2010). Their results use the standard notion of a substructure. However, they only consider the case of partial rationalizability, while we allow for a range of possibilities, including exact rationalizability as well as sub- and supra-semirationality. Our notion of a solution is more general, as we place no restriction on the first-order sentence ϕ defining the solution. Also, we use a stronger notion of falsifiability, *FIT*-ness.

3 Eliminability and complexity

In this section, we review the notion of eliminability of theoretical terms, and offer a new perspective on the meaning of (strong) Ramsey eliminability. In addition, we address the question of eliminability of theoretical terms, i.e., preference relations, for choice theories.

3.1 Eliminability

Suppose that $\mathcal{T} = \langle O_1, \ldots, O_k, T_1, \ldots, T_m \rangle$ is a relational language, with the O_i describing observable relationships, and the relations T_i being theoretical terms. Suppose $\psi(O_1, \ldots, O_k, T_1, \ldots, T_m)$ is a first-order sentence in $\mathcal{T}$, i.e., a (finitely axiomatizable) $\mathcal{T}$-theory. Ramsey (1931) introduced the idea of rewriting ψ without the theoretical terms using what is now called a *Ramsey sentence*:

$$\exists_2 X_1, \ldots, X_m \psi(O_1, \ldots, O_k, X_1, \ldots, X_m). \tag{9}$$

This sentence is meant as an observationally equivalent way to write ψ, but without the theoretical terms. Instead, second-order existential quantification is introduced, with predicate variables X_i. Rewriting a theory without the theoretical terms is called *Ramsey elimination*. Several attempts have been made to define and to study Ramsey eliminability formally. Hintikka (1998) provides a review of the relevant notions, and follows Sneed (1971) in strengthening the notion of Ramsey eliminability to *strong Ramsey eliminability*: there exists a first-order sentence $\rho(O_1, \ldots, O_k)$ that is equivalent to (9). The meaning and significance of strong Ramsey eliminability have been viewed differently by different authors. As (Hintikka, 1998, p. 294) points out, the mere introduction of a Ramsey sentence does nothing to eliminate the effect that theoretical terms have:

> In order to do their job, theoretical concepts merely need to exist in certain relations to the observational concepts of the theory, but it does not matter how or where they exist. In such a model-theoretical perspective, the Ramsey reduction does not change anything at all. It merely means that the left hand (predicate constants) lends money (condition-imposing power) to the right hand (second-order quantifiers).

In fact, Hintikka (1998) argues (making use of *independence-friendly first-order logic*) that the notion of Ramsey eliminability merely points to the structural role that quantifiers play in a theory.

In this subsection we will propose that various notions of Ramsey eliminability could be based on natural complexity classes. We will point out that well-known results in the field of descriptive complexity theory make it possible to connect the notion of Ramsey eliminability with notions of computational and descriptive complexity, and that these connections add another interpretation to Ramsey eliminability.

Descriptive complexity theory[7] relates the logical language required to describe a property (its *descriptive complexity*) to the computational complexity of checking that property. It is a subfield of finite model theory, and applies to finite structures only—indeed, questions of computational complexity would not make sense for infinite structures.

Assumption 2. In this section, we assume that all structures are finite.

[7] See Immerman (1995) for a short introduction.

Definition 9. Let $\mathcal{T} = \langle O_1, \ldots, O_k, T_1, \ldots, T_m \rangle$ be a relational language, with the O_i describing observable relationships, and the relations T_i being theoretical terms, and let ψ be a $\mathcal{T}$-sentence. Let $\mathcal{O} = \langle O_1, \ldots, O_k \rangle$ be the relational language for observations, and let $STRUC[\mathcal{O}]$ be the set of $\mathcal{O}$-structures. We define[8] the *falsifiability query* $\mathbf{Q}(\psi) : STRUC[\mathcal{O}] \to \{0, 1\}$ as: for all $\mathcal{O}$-structures A,

$$\mathbf{Q}(A) = 1 \iff \text{there exists a } \mathcal{T}\text{-structure } A' \text{ that extends } A \text{ to satisfy } \psi. \quad (10)$$

Fundamental results in descriptive complexity theory relate the *computational* complexity of the query $\mathbf{Q}$ to the *descriptive* complexity of the logical language necessary to define the set of $\mathcal{O}$-structures A such that $\mathbf{Q}(A) = 1$. Fagin's Theorem (Fagin, 1973) was the first result establishing a close correspondence between computational and descriptive complexity. It shows that the class *NP* (non-deterministically polynomial queries) is equal to the class $SO\exists$ (queries that can be defined using existential second-order logic).[9] The following result, stated in Chambers et al (2010), follows immediately from Fagin's Theorem. We use the notation as in Definition 9 above.

Theorem 2. *Suppose ψ is a first-order $\mathcal{T}$-structure. Then the computational problem of determining whether an $\mathcal{O}$-structure is extendable to satisfy ψ is in the class NP.*

Proof. The Ramsey sentence of ψ defines the query in the theorem. The Ramsey sentence is in $SO\exists$, so by (one direction of) Fagin's Theorem, the query is in *NP*. $\square$

The falsifiability query being in *NP* means that it can be *verified* in polynomial time that a particular structure is extension of the observation structure satisfying ψ. Problems that are not polynomial are usually considered to be intractable, and so it would be interesting to characterize those theories ψ for which not only is *verification* polynomial, but actually determining *whether* an $\mathcal{O}$-structure is extendable to ψ is polynomial.

Just as the theorem of Fagin (1973) shows that the computational complexity class *NP* is equal to a natural descriptive complexity class ($SO\exists$), the Immerman-Vardi Theorem (Immerman (1982); Vardi (1982); stated as Theorem 4.10 in Immerman (1999)) shows that the computational complexity class *P* is equal to the descriptive complexity class *FO(LFP)*, the class of queries that can be stated using first-order logic extended by the *least fixed point* (*LFP*) operator. The *LFP* operator

[8]Our *falsifiability query* is a boolean query—see Immerman (1999) for more on boolean and more general types of queries.

[9]We refer the interested reader to the authoritative book by Immerman (1999) for an in-depth exposition.

adds the power of inductive definitions to first-order logic. To define it, let $\mathcal{L}$ be a relational language, A an $\mathcal{L}$-structure, R a new relation of arity k, and $\phi(R, x_1, \ldots, x_k)$ a first-order formula in which R occurs only positively. The operator F_ϕ, defined by

$$F_\phi : R \mapsto \{(a_1, \ldots, a_k) : (A, R) \vDash \phi(R, a_1, \ldots, a_k)\}, \tag{11}$$

maps the relation R to a new relation of arity k. This operator has a fixed point R^*, i.e., $F_\phi(R^*) = R^*$, and a least (by containment) fixed point.[10] The *LFP* operator assigns this least fixed point to the formula ϕ in structure A. One important example of an *LFP* operator is the *TC* operator, which maps a binary relation to its transitive closure. Thus the descriptive complexity class *FO(TC)* is first-order logic extended by the transitive closure operator.

Though the Immerman-Vardi theorem holds only for *ordered* structures, the direction we use below does not rely on the ordering assumption.[11] We state the following immediate corollary of the Immerman-Vardi Theorem:

Theorem 3. *Suppose ψ is a first-order $\mathcal{T}$-sentence. If the Ramsey sentence for ψ can be rewritten in first-order logic, then the computational problem of determining whether an $\mathcal{O}$-structure is extendable to ψ is polynomial. Thus, strong Ramsey eliminability implies the computational tractability of falsification.*

Remark 6. It would be desirable to strengthen this result by identifying a natural complexity class that is *equivalent* to strong Ramsey eliminability. However, a theorem by van Benthem shows that "logic has no syntactic criterion available for establishing if a theory has this property." (Benthem, 1978, Theorem 4.8) Thus it is unlikely that there would be a natural complexity class characterizing strong Ramsey eliminability.

For finite structures, theorems 2 and 3 give meaning to the notion of (strong) Ramsey eliminability in terms of the tractability of the computational problem of falsification. Though determining falsifiability is, in principle, always possible for finite structures, the computational tractability of determining whether a particular structure is extendable to a theory is often of great interest. Strong Ramsey eliminability is a non-trivial requirement that guarantees the computational tractability of falsification.

Furthermore, the connection with descriptive complexity suggest a refinement of the notion of Ramsey eliminability. As we noted above, the elimination of theoretical terms via a Ramsey sentence is, by itself, too weak a requirement. The motivation

[10]For these results and further details, see (Grädel, 2007, p. 153).

[11]See the discussion on p. 61 of Immerman (1999).

for strengthening it is to express the property that a theory may be described using its observational terms only, and in a less complex logic than second-order existential logic. First-order logic is certainly less complex, but there are other descriptive complexity classes of intermediate complexity. The Immerman-Vardi theorem suggests that *FO(LFP)*, first-order logic extended by inductive definitions, is of particular interest because of its connection with computational tractability, but other complexity classes may be of interest as well.

Definition 10. Suppose $\psi(O_1, \ldots, O_k, T_1, \ldots, T_m)$ is a first-order sentence in $\mathcal{T}$, i.e., a (finitely axiomatizable) $\mathcal{T}$-theory. If the *Ramsey sentence*

$$\exists_2 X_1, \ldots, X_m \psi(O_1, \ldots, O_k, X_1, \ldots, X_m)$$

of ψ is equivalent to a sentence $\rho(O_1, \ldots, O_k)$ in the descriptive complexity class Ξ, we say that ψ is Ξ *Ramsey eliminable.*

Using this terminology, strong Ramsey eliminability is simply *FO* Ramsey eliminability. We may now state a stronger version of the theorem above:

Theorem 3'. *Suppose ψ is a $\mathcal{T}$-sentence. If ψ is $FO(LFP)$ Ramsey eliminable, then the computational problem of determining whether an $\mathcal{O}$-structure is extendable to ψ is polynomial. Thus, FO(LFP) Ramsey eliminability implies the computational tractability of falsification.*

The Church-Turing thesis makes it possible to study computational complexity independent of a particular model of computation. Results on descriptive complexity theory establish that computational complexity classes correspond naturally to descriptive complexity classes. As the results above demonstrate, this opens the possibility of using formal logic to study not only the falsifiability of theories, but also the computational tractability of actually determining whether a set of observations falsifies a theory. The new notions of Ramsey eliminability we introduce are relevant only for finite structures, but finiteness is often a natural assumption, especially in choice theories. In the next section we show that, in the case of (finite) choice theory, the distinction of tractability versus non-tractability of falsification is a non-trivial one. In particular, falsification is tractable for the theory of individual preference maximization, but it is not for the theory of Nash equilibrium.

3.2 Eliminability in choice theories

In choice theories, the theoretical terms are the preference relations of the decision makers. The Ramsey sentence for a choice theory therefore states that "there exist preferences such that the alternatives observed as chosen are solutions (for sub-semirationality), or that the alternatives not observed as chosen are not solutions (for supra-semirationality), or both (for exact rationality)." Identifying the required preferences is referred to as "rationalizing the observed choices," or, when the solution is Nash equilibrium, "Nash rationalizing the observed choices."

3.2.1 Individual preference maximization

The well-known theorems of revealed preference theory (see, e.g., Richter, 1966) can be viewed as showing that various kinds of theories of individual preference maximization are *FO(TC)* Ramsey eliminable. However, not even the weakest theory of individual preference maximization, partial rationalizability, satisfies *FO* (or strong) Ramsey eliminability. The results in this subsection use the notation and assumptions established in section 2 (see (5) and (6)). To avoid the trivial cases of rationalization by complete indifference, we focus on the strict preference versions of these theories.

Theorem 4.

(i) *The theories of individual strict preference maximization are not FO Ramsey eliminable, whether exact rationality or sub-semirationality (or "partial" rationality) is required. That is,* $\Psi^1(\text{str-}yR_1^2z)$ *and* $\Psi^1_{sub}(\text{str-}yR_1^2z)$ *are not FO Ramsey eliminable.*

(ii) *The theories of individual strict preference maximization are FO(TC) Ramsey eliminable, when exact rationality or sub-semirationality is required. That is,* $\Psi^1(\text{str-}yR_1^2z)$ *and* $\Psi^1_{sub}(\text{str-}yR_1^2z)$ *are FO(TC) Ramsey eliminable.*

Proof. (i) This result follows easily from the fact that acyclicity is not first-order expressible (see Immerman, 1999, Proposition 6.24). *(ii)* The Strong Axiom of Revealed Preference (SARP) is the *FO(TC)* sentence that is equivalent to the Ramsey sentence of these theories (see Houthakker, 1950; Richter, 1966). □

It follows from Theorems 3' and 4 that determining whether an observational structure is exactly or partially strict preference rationalizable is a polynomial problem (see also Galambos, 2011).

3.2.2 Nash equilibrium

The literature examining questions of testability in game theory is fairly recent (see, e.g., Carvajal et al, 2004; Yanovskaya, 1980; Sprumont, 2000). Galambos (2011) showed that the problem of Nash rationalizability is *NP*-complete. This immediately implies the following result, proved in (Galambos, 2011) and stated here using the new notions we introduced. Recall that the theory of Nash equilibrium is $\Psi^n(\varphi_{\text{Nash}})$, with φ_{Nash} the first-order sentence in (4).

Theorem 5. *The theory of Nash equilibrium is not FO(LFP) Ramsey eliminable, unless $P = NP$.*[12]

This contrasts with the *FO(TC)* Ramsey eliminability of the theory of individual preference maximization (Theorem 4 above). Even though both theories are *FIT* (Theorem 1 above), using more discriminating notions of Ramsey eliminability, we see that the degree of testability of the dominant individual and collective choice theories is different. This is significant because these more refined notions of Ramsey eliminability reflect the practical (computational) difficulty of determining whether a particular set of observations falsifies a theory.

Our result differs from that in Chambers et al (2010), who are not able to differentiate the degree of testability of individual preference maximization and Nash equilibrium. The main reason for this difference is that they consider only *partial* Nash rationalizability, that is, the theory $\Psi^n_{sub}(\varphi_{\text{Nash}})$. It is not difficult to see that this partial rationalizability version of the theory of Nash equilibrium is, in fact, *FO(TC)* Ramsey eliminable, and the *FO(TC)* equivalent of its Ramsey sentence is a straightforward generalization of the Strong Axiom of Revealed Preference (see Galambos, 2004).

4 Conclusion

At least since the 1930s, questions of falsifiability have been studied in the economics literature. We show that choice theories are falsifiable in a strong sense, as defined by the notion of *FIT*-ness. This notion has sometimes been found to be too demanding in the natural sciences. With an appropriate adaptation of the notion of a substructure to reflect the structure of observations, we find that *FIT*-ness is not too strong a requirement for choice theories.

More generally, we relate notions of complexity to the question of eliminability of theoretical terms. We propose that degrees of testability can be studied using descriptive complexity classes between second-order and first-order logic. Fundamental

[12]The qualification in the theorem is a reminder that the $P \stackrel{?}{=} NP$ is one of the most important open questions in mathematics today. It is widely believed that $P \subsetneq NP$, and many results in complexity theory are qualified in this way.

results in descriptive complexity theory make it possible to interpret various types of eliminability in terms of the computational tractability of falsification. When the restriction to finite structures is natural, these ideas provide further tools in the study of testability of scientific theories.

Through the example of the most fundamental choice theories, individual preference maximization and Nash equilibrium, we show how the notions of Ramsey eliminability we introduce can be used to differentiate the degree of testability of these theories.

References

Backhouse R (ed) (1994) New directions in economic methodology. Economics as social theory, Routledge, URL http://books.google.com/books?id=8BVzbC_xVpAC

Benthem JFAKv (1978) Ramsey eliminability. Studia Logica: An International Journal for Symbolic Logic 37(4):pp. 321–336, URL http://www.jstor.org/stable/20014913

Boland LA (1981) On the futility of criticizing the neoclassical maximization hypothesis. The American Economic Review 71(5):pp. 1031–1036, URL http://www.jstor.org/stable/1803485

Carvajal A, Ray I, Snyder S (2004) Equilibrium behavior in markets and games: testable restrictions and identification. Journal of Mathematical Economics 40(1-2):1–40

Chambers CP, Echenique F, Shmaya E (2010) General revealed preference theory. Social Science Working Paper 1332, California Institute of Technology

Fagin R (1973) Contributions to the model theory of finite structures. PhD thesis, U.C. Berkeley

Galambos A (2004) Revealed preference in game theory. PhD thesis, University of Minnesota

Galambos A (2011) The complexity of Nash rationalizability. Tech. rep., Lawrence University

Grädel E (2007) Finite model theory and descriptive complexity. In: Finite Model Theory and Its Applications, Springer, pp 125–230, URL http://www.logic.rwth-aachen.de/pub/graedel/FMTbook-Chapter3.pdf

Hausman D (1992) Essays on philosophy and economic methodology. Cambridge University Press, URL http://books.google.com/books?id=pM57-DJ2yDcC

Hintikka J (1998) Ramsey sentences and the meaning of quantifiers. Philosophy of Science 65(2):pp. 289–305, URL http://www.jstor.org/stable/188262

Houthakker HS (1950) Revealed preference and the utility function. Economica NS 17(66):159–174

Hutchison T (1938) The significance and basic postulates of economic theory. Macmillan, URL http://books.google.com/books?id=N8lfAAAAIAAJ

Immerman N (1982) Relational queries computable in polynomial time. In: 14th Symposium on Theory of Computation, ACM, pp 147–152

Immerman N (1995) Descriptive complexity: a logician's approach to computation. Notices of the American Mathematical Society 42

Immerman N (1999) Descriptive Complexity. Graduate Texts in Computer Science, Springer

Matzkin RL, Richter MK (1991) Testing strictly concave rationality. Journal of Economic Theory 53:287–303

Mongin P (1986) Are "all-and-some" statements falsifiable after all?: The example of utility theory. Economics and Philosophy 2(2):185–195

Popper K (1962) Conjectures and refutations: the growth of scientific knowledge. Routledge & Kegan Paul, URL http://books.google.com/books?id=ZHRAAQAAIAAJ

Ramsey F (1931) Theories. In: Braithwaite R (ed) The foundations of mathematics and other logical essays, no. v. 5 in The international library of philosophy: Philosophy of logic and mathematics, Routledge and Kegan Paul

Richter MK (1966) Revealed preference theory. Econometrica 34:635–645

Rynasiewicz R (1983) Falsifiability and the semantic eliminability of theoretical languages. The British Journal for the Philosophy of Science 34(3):225–241

Samuelson PA (1953) Consumption theorems in terms of overcompensation rather than indifference comparisons. Economica 20(77):pp. 1–9, URL http://www.jstor.org/stable/2550984

Simon HA, Groen GJ (1973) Ramsey eliminability and the testability of scientific theories. The British Journal for the Philosophy of Science 24(4):pp. 367–380, URL http://www.jstor.org/stable/686617

Sneed J (1971) The logical structure of mathematical physics. Synthese library, Reidel

Sprumont Y (2000) On the testable implications of collective choice theories. Journal of Economic Theory 93:205–232

Vardi MY (1982) Complexity of relational query languages. In: 14th Symposium on Theory of Computation, ACM, pp 137–146

Yanovskaya E (1980) Revealed preference in non-cooperative games. Mathematical Methods in the Social Sciences 24(13), in Russian