GENERAL REVEALED PREFERENCE THEORY

CHRISTOPHER P. CHAMBERS, FEDERICO ECHENIQUE, AND ERAN SHMAYA

Abstract. We generalize the standard revealed-preference exercise in economics, and prove a sufficient condition under which the revealed-preference formulation of an economic theory has universal and effectively testable implications. We not only generalize and "explain" classical revealed-preference theory, but we also obtain applications to the theory of group preference and Nash equilibrium.

1. Introduction

The notion of revealed preference was introduced by Paul Samuelson (1938) in his investigation of the empirical content of the theory that consumers maximize utility. In this paper we generalize the standard revealed-preference exercise in economics: We provide a general framework that captures Samuelson's ideas as a special case; but which applies to many other economic models.

Samuelson formulated the theory of the consumer as a statement about observable data. He attempted to characterize the data sets that are consistent with the *existence of some* utility function. The answer, provided by Houthakker (1950), is that the data sets that are

---

Chambers is affiliated with the Department of Economics, University of California, San Diego. Echenique is affiliated with the Division of the Humanities and Social Sciences, California Institute of Technology, Pasadena CA 91125. Shmaya is affiliated with the Kellogg School of Management at Northwestern University, 2001 Sheridan Road, Evanston, IL 60208-2001. Chambers and Echenique acknowledge support from the NSF through grant SES-0751980. Emails: `cpchambers@ucsd.edu` (Chambers), `fede@caltech.edu` (Echenique), and `e-shmaya@kellogg.northwestern.edu` (Shmaya). We are grateful to seminar audiences at Arizona State University and Brown University. We are especially grateful to Kim Border and Philippe Mongin for many detailed comments and suggestions.

consistent with utility maximization are those that satisfy the strong axiom of revealed preference (SARP). It is crucial to understand the importance of SARP, because it uncovers a basic issue behind any revealed preference exercise, and this issue is at the heart of our paper.

We often hear that SARP contains the testable implications of the theory of utility maximization, but the theory is in principle testable even without access to a result like Samuelson's and Houthakker's. As an empirical statement, the theory of the consumer comprises the data sets that *some* utility function can explain. This statement describes the empirical content of the theory of the consumer just as SARP does. Importantly, though, it is useless as a test; because to test the theory of the consumer we would need to check every possible utility function and see if they could have generated the data—a completely impractical test. In contrast, SARP provides an "effective" test. Given any data set falsifying the theory of the consumer, we can *in finitely many steps* determine that it violates SARP.

An example from Popper (1959) illustrates our point. Suppose that theory E says "There is a black swan;" while theory U says "All swans are white." Theory E is not falsifiable because no matter how many finite data sets of non-black swans we find, it is still possible that there is somewhere a black swan. Theory U is falsifiable because the observation of a single non-white swan contradicts the theory. Note the similarity between E and the first formulation of utility maximization, "There is some utility function that explains the data." Such a formulation is *existential*, and it does not lead to a test of the theory for the same reason that theory E is not testable. In contrast, theory U and SARP are *universal* statements. They are testable because they do not involve any unobservable free parameters.

Our main result provides the existence of SARP-like universal tests for any revealed-preference exercise satisfying the assumption that the theory would be testable if only its unobservable terms were observable. More specifically, we prove that if the formulation of a theory is universal when we assume that its unobservable terms are observable, then there is a test purely based on observables; this test is universal:

it is a universal axiomatization of the data sets that comply with the theory, analogously to SARP. When the formulation over unobservables is effective (which it is in all applications we can think of), then our test is also effective, in the sense that a falsifying data set can be certified to be a falsification in finitely many steps.[1]

We present applications of our result to the problem of testing for Nash equilibrium, and to other theories of group decision making. Our result is easy to apply: we need to write the theory of Nash equilibrium, for example, using purely universal axioms (axioms that only involve "for all" quantifiers). Some axioms would make statements about preferences: for example if we are talking about two-player games we can write "if (x, y) is observed as an outcome, and $z \neq x$ is a strategy for player 1, then player one prefers (x, y) to (z, y)." Such an axiomatization would be a test of the theory in the unrealistic case in which preferences are observable. Our main result implies then that there is a test (a universal axiomatization) that is purely based on observables, and that this test certifies a falsification in finitely many steps.

Our results are similar in spirit to the applications of "quantifier elimination" in revealed preference theory, see Brown and Matzkin (1996) and Brown and Kubler (2008). Brown and Matzkin show that the revealed preference formulation of Walrasian equilibrium theory leads to testable implications on economy-wide data. That is, there exist preferences which could generate the data if and only if a collection of universal statements on data are satisfied. Quantifier elimination and our results have in common a kind of "projection" of a theory onto purely observable terms. Our results use completely different techniques, however. As we explain below, we use a result due to Alfred Tarski on universal axiomatizations in model theory. The applications

[1]In fact, the problem of falsification is decidable; in the sense that there is an algorithm that determines if a finite instance of data is compatible with the theory. A direct application of an important theorem from computer science (Fagin's theorem) implies that the theories in our framework lie in the complexity class NP. This means that, given a structure which is in our theory, there is a polynomial-time way of *verifying* that it is actually in our theory.

we present in this paper cannot be analyzed by means of quantifier elimination.[2]

We proceed to summarize the applications of our main result.

Group preference and group choice. Consumption theory involves one unobserved preference relation. Our general result allows there to be multiple unobserved relations.

Two branches of the revealed preference literature focus on the empirical content of group choices: one motivated by social choice considerations, the other by game-theoretic considerations.

One branch is devoted to understanding the empirical content of collections of individuals who have conflicting motivations. We observe group preferences, but we do not observe the preferences of the underlying members of the group, nor do we observe the method by which they aggregate their preferences. We want to test the joint hypothesis that the individuals are rational, and use some particular social choice rule in aggregating their preferences. Exercises of this type date back at least to Dushnik and Miller (1941), which can be interpreted as axiomatizing those binary relations which could be written as the Pareto relation for two individuals.

We establish that for a broad class of preference aggregation rules, namely those which are neutral and satisfy independence of irrelevant alternatives, the hypothesis that preference is derived from an aggregate of individual rational preferences is universally axiomatizable and hence falsifiable in principle.

Another recent branch of the revealed preference literature focuses on the empirical content of group choice functions in games. The approach is as follows: a set of players is given, and a set of strategies is given. For any nonempty set of strategies, for each agent, it is imagined that a joint choice is observed from the game form derived from those sets. One can ask whether or not the observed choices can be rationalized as Nash equilibria, or as the Pareto optimal joint choices for some

[2]At least not by the quantifier elimination of real-closed fields, as in Brown and Matzkin (1996). We deal with theories that are not necessarily formulated in the language of real closed fields.

set of preferences. Examples of such papers include Peleg and Tijs (1996), Sprumont (2000), Xu and Zhou (2007), Galambos (2009), and Lee (2009). There are older studies of the same questions, also about group choice, but using other solution concepts: see Wilson (1970), Plott (1974), and Ledyard (1986).

Here, we work with a generalized notion of equilibrium which incorporates Nash equilibrium, strong Nash equilibrium (Aumann, 1960), and Pareto optimality as special cases. We show that the theory hypothesizing that there exist *strict* preferences rationalizing the observed choices is always universally axiomatizable, and hence falsifiable in principle.

Section 2 lays down the formal structure of the model, and presents the main result (Theorem 5). Section 3 presents the application to group choice and preference aggregation. Section 4 discusses the application to group choice behavior. Section 5 establishes the complexity result (Fagin's theorem). Section 6 talks about our large sample result, while Section 7 discusses previous literature and concludes. The Appendix collects details about model theory, in order to keep our presentation self-contained.

2. Main Results

2.1. **Preliminary definitions.** In our setup, a test is an axiomatization of possible data; so we are going to theorize about axiomatizations. To this end we borrow ideas and results from the field of model theory in mathematics. We proceed to give some standard definitions from model theory. Readers with at least a minimal exposure to model theory or mathematical logic will want to skip this section.

We first must specify our language $\mathcal{L}$. The language is a primitive and specifies the *syntax*, or the things we can say.

1. **Definition.** A *relational first-order language* $\mathcal{L}$ is given by a set $\mathcal{R}$ of predicate symbols and a positive integer n_R for every $R \in \mathcal{R}$.

The symbol $R \in \mathcal{R}$ is meant to denote a n_R-ary relation. Note that we focus here on language without constant and function symbols.

(See also Remark 6 below). When R is a binary relation, we use the notation $R(x, y)$ instead of $x\, R\, y$. When R is a ternary relation, we write $R(x, y, z)$, and so on.

For example if we wish to talk about preference, we may use a language with a single binary symbol R. We can then write axioms, such as $\forall x R(x, x)$, or $\forall x \exists y R(y, x)$, and make sense of when a set X with a specific binary relation on X satisfy these axiom. Issues that deal with the form of axioms are issues of *syntax*; while specific sets and relations raise *semantical* issues.

The semantics are specified by concrete mathematical objects, called *structures.* Structures provide the appropriate framework for interpreting our syntax. Specifically, a structure is a universe of possible objects, and an interpretation of the elements of language in that universe.

2. **Definition.** An $\mathcal{L}$-*structure* $\mathcal{M}$ is given by a nonempty set M called the *domain* of $\mathcal{M}$, and, for every predicate symbol $R \in \mathcal{R}$, a set $R^{\mathcal{M}} \subseteq M^{n_R}$.

When the language $\mathcal{L}$ is understood, we refer to an $\mathcal{L}$-structure simply as a *structure*. The elements $R^{\mathcal{M}}$ and are called *interpretations* of the corresponding symbols in the language $\mathcal{L}$.[3] For example, when $\mathcal{L} = \langle \succeq \rangle$ has a single binary relation, then one possible structure is $(\mathbf{R}, \geq)$; the structure of the real numbers with the usual greater-than binary relation.

3. **Definition.** Suppose that $\mathcal{M}$ and $\mathcal{N}$ are $\mathcal{L}$-structures with universes M and N respectively. An $\mathcal{L}$-*embedding* $\eta : \mathcal{M} \to \mathcal{N}$ is a one-to-one map $\eta : M \to N$ that preserves the interpretations of all predicate symbols, i.e., such that

$$R^{\mathcal{M}}(a_1, ..., a_{m_R}) \leftrightarrow R^{\mathcal{N}}(\eta(a_1), ..., \eta(a_{m_R}))$$

for all $R \in \mathcal{R}$ and $a_1, ..., a_{m_R} \in M$.

As a notational convention, we write $R^{\mathcal{M}}(a_1, ..., a_{m_R})$ to mean $(a_1, ..., a_{m_R}) \in R^{\mathcal{M}}$.

[3]We make no cardinality restrictions on structures.

4. **Definition.** An *isomorphism* is a bijective $\mathcal{L}$-embedding.

Appendix A gives definitions of sentence, and of the validity of a sentence in a structure. These notions correspond closely to their conventional meaning in English.

2.2. **The Model.** For a given first order language $\mathcal{G}$, a $\mathcal{G}$-theory is a class of structures for that language which is closed under isomorphism. We say that a theory T is *axiomatized* by a collection of sentences Σ if T consists exactly of the structures for which each sentence in Σ is valid. Given two theories, T and T', where $T \subseteq T'$, we say that T is *axiomatized* by a collection of sentences Σ *with respect to* T' if T consists of exactly those structures in T' for which each sentence in Σ is valid.

Let $\mathcal{F} = \langle R_1, \ldots, R_N \rangle$ and

$$\mathcal{L} = \langle R_1, \ldots, R_N, Q_1, \ldots, Q_K \rangle,$$

be languages, where all the R_n and Q_k are predicate symbols. Note that $\mathcal{F} \subseteq \mathcal{L}$.

The languages $\mathcal{F}$ and $\mathcal{L}$ are our way of capturing the difference between observed and unobserved terms. We mean the relations R_n to be observable in the data, while the relations Q_k are unobservable. In our applications below, we choose $\mathcal{F}$ and $\mathcal{L}$ with this interpretation in mind. For example, for the standard revealed preference theory of individual rational choice, $\mathcal{F} = \langle R \rangle$ and $\mathcal{L} = \langle R, Q \rangle$, where R is meant to represent the revealed preference relation present in the data, and Q is meant to be a theoretical preference relation, on which we impose some rationality axioms (see Section 2.4 below).

Let T be an $\mathcal{L}$-theory. Define $F(T)$ to be the class of $\mathcal{F}$-structures $(X^*, R_1^*, ..., R_N^*)$ for which there exist relations $Q_1^*, ..., Q_K^*$ such that $(X^*, R_1^*, ..., R_N^*, Q_1^*, ..., Q_K^*) \in T$. That is, $F(T)$ is the *projection* of T onto the language $\mathcal{F}$.[4]

[4]If T has a finite first-order axiomatization, then $F(T)$ is an example of an **existential second-order** theory for language $\mathcal{F}$, in that it allows existential quantification over predicates. That is, if σ is a first-order $\mathcal{L}$-axiom axiomatizing theory T, then

Recall that a *universal* axiom is an axiom of the form $\forall x_1 ... \forall x_n \varphi(x_1, ..., x_n)$, where $\varphi(x_1, ..., x_n)$ is quantifier free. A universal axiomatization is an axiomatization that consists entirely of universal axioms.

Given a collection of $\mathcal{L}$-sentences Σ, the collection of *$\mathcal{F}$-consequences of* Σ is the collection of all logical implications of $\mathcal{L}$ involving only predicates from $\mathcal{F}$.

5. **Theorem.** *If T has a universal $\mathcal{L}$-axiomatization, then $F(T)$ has a universal $\mathcal{F}$-axiomatization. Moreover, if Σ is a universal axiomatization of T, then the collection of universal $\mathcal{F}$-consequences of Σ is an axiomatization of $F(T)$.*

6. *Remark.* For readers who are familiar with model theory, we remark that Theorem 5 remains valid when the language $\mathcal{F}$ includes also constant and function symbols. However, the unobservable symbols $Q_1, \ldots, Q_K$ must be only predicate symbols.

7. *Remark.* In Section 7.1 we present an example of how $F(T)$ can fail to have an axiomatization when T is not under the hypothesis of Theorem 5.

Proof. We first establish that $F(T)$ is universally axiomatizable, using a theorem of Tarski (1954). We then show what this axiomatization should be.

We want to verify the three conditions of Theorem 1.2 in Tarski (1954). Specifically, conditions (i), (ii), and (iii') in Tarski's paper. To this end, we need to show that $F(T)$ is closed under isomorphism and substructure. Lastly, we need to show that for any totally ordered set Θ and indexed collection of models $\mathcal{M}^\theta \in F(T)$ where if $\theta < \theta'$, then $\mathcal{M}^\theta$ is a substructure of $\mathcal{M}^{\theta'}$, there is $\mathcal{M}^* \in F(T)$ for which each $\mathcal{M}^\theta$ is a substructure of $\mathcal{M}^*$.

$F(T)$ is that $\mathcal{F}$-theory axiomatized by

$$\exists Q_1 ... \exists Q_K \sigma.$$

The first two conditions (closure under substructures and isomorphism) follow because T satisfies those conditions as it is a universal theory. We prove the third condition.

Let Σ be a universal axiomatization of T.

Let $\mathcal{M}^\theta = (X^\theta, R_1^\theta, ..., R_N^\theta)$, $\theta \in \Theta$, be a monotone class of structures of $F(T)$. That is, Θ is totally ordered, and $\mathcal{M}^\theta$ is a substructure of $\mathcal{M}^{\theta'}$ whenever $\theta < \theta'$. Let $\mathcal{M}^* = (X^*, R_1^*, \ldots, R_N^*)$ be defined so that $X^* = \bigcup_\theta X^\theta$ and $R_k^* = \bigcup_\theta R_k^\theta$ for $k = 1, \ldots, N$.

For each θ, let W^θ be the set of lists of relations $(Q_1^*, \ldots, Q_K^*)$ on X^* such that

$$(X^\theta, R_1^\theta, \ldots, R_N^\theta, Q_1^*|_{X^\theta}, \ldots, Q_K^*|_{X^\theta})$$

is a model of Σ. Note that $W^\theta \neq \varnothing$ as $\mathcal{M}^\theta \in F(T)$.

We claim that if $\theta < \theta'$, then $W^{\theta'} \subseteq W^\theta$.

Let $(Q_1^*, \ldots, Q_K^*) \in W^{\theta'}$ and let $\forall x_1 ... \forall x_M \varphi(x_1, \ldots, x_M) \in \Sigma$. Then by assumption, $(X^{\theta'}, R_1^{\theta'} \ldots, R_N^{\theta'}, Q_1^*|_{X^{\theta'}}, \ldots, Q_K^*|_{X^{\theta'}})$ is a model of $\forall x_1 ... \forall x_M \varphi(x_1, ..., x_M)$, so $\varphi(x_1^*, \ldots, x_M^*)$ is valid for any $\{x_1^*, \ldots, x_M^*\} \subseteq X^{\theta'}$; in particular, it is valid for any $\{x_1^*, \ldots, x_M^*\} \subseteq X^\theta$. Since for all $i = 1, \ldots, N$, $R_i^\theta = R_i^{\theta'} \cap X^\theta$, $\forall x_1 ... \forall x_M \varphi(x_1, ..., x_M)$ is valid in $(X^\theta, R_1^\theta, \ldots, R_N^\theta, Q_1^*|_{X^\theta}, \ldots, Q_K^*|_{X^\theta})$. As φ was arbitrary, $(Q_1^*, \ldots, Q_K^*) \in W^\theta$.

Note that if Q is a k-ary relation on X^*, it is a subset of X^k. To this end, regard W^θ as a subset of

$$\mathcal{B} = \{0,1\}^{\Pi_1 X^*} \times \cdots \times \{0,1\}^{\Pi_K X^*},$$

where $\Pi_k X^*$ stands for the product $X^* \times \ldots, \times X^*$, as many times as the order of the predicate Q_k. Note that $\mathcal{B}$, endowed with the product topology, is compact.

We claim that W^θ, viewed as a subset of $\mathcal{B}$, is closed. To see this, let $(Q_1^\lambda, \ldots, Q_K^\lambda)$ be a net in W^θ, converging to $(Q_1^*, \ldots, Q_K^*)$. Let $\forall x_1 ... \forall x_M \varphi(x_1, \ldots, x_M)$ be a formula in Σ and let $\{x_1^*, \ldots, x_M^*\} \subseteq X^\theta$. Then by definition of product topology convergence, there exists $\overline{\lambda}$ such that if $\overline{\lambda} < \lambda$, then for $k = 1, \ldots, K$, $Q_k^\lambda(x_1^* \ldots, x_M^*)$ if and only if

$Q_k^*(x_1^* \ldots, x_M^*)$. Then, since

$$w = \langle X^\theta, R_1^\theta, \ldots, R_N^\theta, Q_1^\lambda|_{X^\theta}, \ldots, Q_K^\lambda|_{X^\theta} \rangle$$

is a model of $\forall x_1 ... \forall x_M \varphi(x_1, ..., x_M)$, $\varphi(x_1, ..., x_M)$ is valid at $x_1^* \ldots, x_M^*$ for the interpretation of the predicate symbols in w. So $Q_k^\lambda(x_1^* \ldots, x_M^*)$ if and only if $Q_k^*(x_1^* \ldots, x_M^*)$ implies that φ is valid at $x_1^* \ldots, x_M^*$ for the interpretation of the predicate symbols in $(R_1^\theta, \ldots, R_N^\theta, Q_1^*|_{X^\theta}, \ldots, Q_K^*|_{X^\theta})$. Since $(x_1^*, ..., x_M^*)$ was arbitrary, conclude that $(Q_1^*, \ldots, Q_K^*) \in W^\theta$.

The collection W^θ, $\theta \in \Theta$, is thus a nested collection of closed sets in a compact space, so it has nonempty intersection. Conclude that there exists $(Q_1^*, \ldots, Q_K^*) \in \bigcap_{\theta \in \Theta} W^\theta$. We claim that $u^* = (X^*, R_1^*, \ldots, R_N^*, Q_1^*, \ldots, Q_K^*)$ is a model of Σ. Let $\forall x_1 ... \forall x_m \varphi(x_1, \ldots, x_M) \in \Sigma$ and $\{x_1^* \ldots, x_M^*\} \subseteq X^*$. By definition of u^*, there is $\theta \in \Theta$ such that, for $n = 1, \ldots, N$, $R_n^*(x_1^* \ldots, x_M^*)$ if and only if $R_n^\theta(x_1^* \ldots, x_M^*)$, and such that $\{x_1^* \ldots, x_M^*\} \subseteq X^\theta$. Since $(Q_1^*, \ldots, Q_K^*) \in W^\theta$, $\varphi(x_1^*, \ldots, x_M^*)$ is valid under the interpretation $(R_1^\theta, \ldots, R_N^\theta, Q_1^*|_{X^\theta}, \ldots, Q_K^*|_{X^\theta})$. Hence, the fact that $R_n^*(x_1^* \ldots, x_M^*)$ if and only if $R_n^\theta(x_1^* \ldots, x_M^*)$ implies that $\varphi(x_1^*, \ldots, x_M^*)$ is valid under the interpretation $(R_1^*, \ldots, R_N^*, Q_1^*, \ldots, Q_K^*)$. Conclude that $u^* \in T$. As $F(\{u^*\}) = \mathcal{M}^*$, we obtain $\mathcal{M}^* \in F(T)$, establishing the third condition.

Lastly, we now know that $F(T)$ has a universal axiomatization. Obviously, any structure $\mathcal{M} \in F(T)$ satisfies all universal $\mathcal{F}$-implications of Σ. Conversely, we need to show that any sentence in the universal axiomatization of $F(T)$ is a universal $\mathcal{F}$-implication of Σ. So suppose there is a sentence φ which is not. In particular then, there exists a structure $\mathcal{M} \in T$ for which φ is not valid. But as φ involves only predicates from $\mathcal{F}$, it therefore follows that φ is also not valid for $F(\mathcal{M})$, a contradiction (as φ is valid for all members of $F(T)$). $\square$

2.3. **Recursive Axiomatization.** In the introduction we talked of how SARP gives an effective test, one that certifies a falsification in finitely many steps. Note that SARP is actually an infinite collection of axioms, but it still provides an effective test. SARP allows for an algorithm that determines when it has been violated.

Here we formalize the notion of effective tests using ideas from computability theory (Rogers (1987), for example, provides an introduction). The notion of an effective test corresponds to the axiomatization in Theorem 5 being recursively enumerable. Intuitively, we show the existence of axiomatizations for which there is a computer program that can enumerate all the axioms: this is the meaning of recursively enumerable. Any violation of the theory would be detected by our program, as it would check each axiom against the data and see if the data comply with the axiom.

Assume that an economist proposes a theory T, with universal axiomatization Σ. Suppose that he observes the elements $a_1, \ldots, a_n$ of some structure $\mathcal{M}$, and the relationship between them. If there exist some universal axiom $\forall x_1 ... \forall x_n \varphi(x_1, \ldots, x_n) \in \Sigma$ such that $\varphi(a_1, \ldots, a_n)$ is not valid in $\mathcal{M}$ then T has been falsified. On the other hand, if none of the axioms is violated, that does not establish the correctness of the theory, since the axioms may still be violated on other elements of $\mathcal{M}$. This idea is the fundamental tenet of Popper's approach – theories can be falsified, but never be proved.[5]

We now extend this idea to the process of checking whether observed data (that is, a collection of elements $(a_1, \ldots, a_n)$ and the relations between them) falsify the theory. To do that, our economist needs a computer program to produce a list of all the axioms in Σ, so that he can go over the axioms and check their validity over the data set. In this case, if the data violate an axiom, we will eventually find the violations. If none of the axioms are violated, then it is possible that the search will never end and we will never know for sure that none of the axioms is violated. Again, theories are not proved, only falsified.

A set Σ of formulas is called *recursively enumerable* (r.e.) if there exists a Turing Machine that enumerates over the elements of Σ in some

[5] This is referred to as the "asymmetry thesis" in Popper's terminology.

order.[6] Thus, the output of the machine is an exhaustive list $\varphi_1, \varphi_2, \ldots$ of all the elements of Σ. It is easy to see that SARP constitutes a r.e. axiomatization.

The following is a simple corollary of Theorem 5.

8. **Corollary.** *If T has a recursively enumerable and universal axiomatization, then so does $F(T)$.*

Proof. Let Σ be a r.e. set of universal formulas that axiomatizes T. It is well known that the set of logical implications of a r.e. set of formulas is itself r.e. Thus, there exists a Turing machine that enumerates all these logical implications $\psi_1, \psi_2, \ldots$. We augment this Turing machine by checking before printing each ψ_i whether it is a universal sentence that contains only observable predicates, and print ψ_i only if it satisfies these requirements. The augmented machine enumerates over all universal logical implications of Σ that contain only observable predicates. By Theorem 5, this set axiomatizes $F(T)$. $\square$

9. *Remark.* A set is *recursive* if both it and its complement are r.e. By an argument of Craig (1953), a r.e. set Σ of universal axioms is equivalent to a recursive set Σ' of axioms.

To understand the importance of Corollary 8, note that Theorem 5 gives us a universal axiomatization, but it will typically not be a finite one (SARP is our running example, see Section 2.4 for details). To be a proper *test* of the theory, in the same sense as SARP, we need recursive enumerability. In all the applications we have in mind, T has a finite axiomatization, so the corollary ensures that our test is r.e.

2.4. **Example: Individual rational choice.** As an illustration of Theorem 5, consider the revealed preference formulation of the theory of individual rational choice. We begin with the simple case in which revealed preference is the primitive, then we describe the case in which a choice function is primitive.

[6] The set of formulas under consideration is always at most countable. This is a standard notion in mathematical logic; it is at the heart of many celebrated results, the most famous being Gödel's theorem.

Consider the language $\mathcal{F} = \langle R, P\rangle$ with two binary predicates; $R(x, y)$ is intended to mean that x is revealed preferred to y, and $P(x, y)$ that x is revealed strictly preferred to y.

We are interested in the theory of all structures (X, R^X, P^X) for which there exists a complete and transitive binary relation $\succeq$ satisfying the axioms

(1) $\forall x \forall y (R(x, y) \rightarrow \succeq (x, y))$
(2) $\forall x \forall y (P(x, y) \rightarrow \succ (x, y))$

The language $\mathcal{F}$ expresses only observables, but the version we have of the theory does not constitute a universal axiomatization. We want to know when we can formulate the theory using only universal axioms that are statements about observables.

We can extend the language $\mathcal{F}$ to a language $\mathcal{L}$, and formulate our theory using only universal axioms. Let $\mathcal{L} = \langle R, P, \succeq, \succ\rangle$. Here, $\succeq$ and $\succ$ are meant to be theoretical, unobservable, relations. Consider the set of $\mathcal{L}$-sentences:

(1) $\forall x \forall y (\succeq (x, y) \lor \succeq (y, x))$
(2) $\forall x \forall y (\succ (x, y) \leftrightarrow (\succeq (x, y) \land \neg \succeq (y, x)))$
(3) $\forall x \forall y \forall z (\succeq (x, y) \land \succeq (y, z)) \rightarrow \succeq (x, z)$
(4) $\forall x \forall y (R(x, y) \rightarrow \succeq (x, y))$
(5) $\forall x \forall y (P(x, y) \rightarrow \succ (x, y))$

Let T be the $\mathcal{L}$-theory axiomatized by this set of sentences.

Now, T is a description of the theory of rational choice, but it assumes that we can access, or observe, the relation $\succeq$. The theory we are really interested in is $F(T)$: the projection of T onto the language $\mathcal{F}$.

Since the axioms (1)-(5) are a universal axiomatization of T, Theorem 5 implies that there is a universal axiomatization of $F(T)$. In addition, one such axiomatization is given by the implications of (1)-(5) that only involve the predicates R and P.

To be concrete, this axiomatization is described by a variant of the strong axiom of revealed preference (sometimes called Suzumura consistency, after Suzumura (1976)). In fact, in first order logic, the strong axiom of revealed preference is a collection of axioms.

The Strong Axiom of Revealed Preference: For every k,

$$\forall x_1 ... \forall x_k \neg \bigwedge_{i=1}^{k} \left(x_i \, Q_i \, x_{(i+1) \mod k} \right)$$

where for all i, $Q_i \in \{R, P\}$, and for at least one $i \in \{1, ..., k\}$, $Q_i = P$.

10. **Proposition.** *If T is axiomatized by (1)-(5), then $F(T)$ is axiomatized by the strong axiom of revealed preference.*

Proof. We offer a sketch, as this type of argument is well-understood. Clearly the strong axiom is valid for $F(T)$. Now suppose that (X, R^X, P^X) is a model of the theory described by the strong axiom. We want to show it is an element of $F(T)$. Let Q denote the transitive closure of $R^X \cup P^X$. Note that if $P^X(x, y)$, then $\neg Q(y, x)$ (this follows by the strong axiom of revealed preference). Consequently, denoting the strict part of Q by P_Q, we obtain $\forall x \forall y R^X(x, y) \to Q(x, y)$ and $\forall x \forall y P^X(x, y) \to P_Q(x, y)$. Now, by a generalization of the Szpilrajn Theorem (see, for example, Suzumura (1976), Theorem 3), Q has an extension to a weak order $\succeq$ with strict part $\succ$, so that $\forall x \forall y Q(x, y) \to \succeq (x, y)$ and $\forall x \forall y P_Q(x, y) \to \succ (x, y)$. Consequently, $\forall x \forall y R^X(x, y) \to \succeq (x, y)$ and $\forall x \forall y P^X(x, y) \to \succ (x, y)$, where $\succeq$ is a weak order and $\succ$ is its strict part. This verifies that $(X, R^X, P^X) \in F(T)$, as $(X, R^X, P^X, \succeq, \succ) \in T$. □

We now turn to a formulation of rational choice theory which uses different primitives, the idea is again to illustrate how our result works. When choice is primitive, the application of our theorem is a bit more involved. This is because we need to be able to describe budget sets, and we require symbols for standard set-theoretic operations. To this end, we define the language of choice as $\mathcal{F} = \langle A, B, \in, C \rangle$. The predicates A and B are unary predicates: $A(x)$ stands for "x is an alternative," $B(x)$ stands for "x is a budget set." The predicates $\in$ and C are binary, $\in$ is the typical set-theoretic predicate, and C is a binary predicate, where $C(x, y)$ means "x is chosen from y."

The theory of choice, T_C, consists of the class of all structures for which there is some global set of alternatives X, and some family of sets $\mathcal{B} \subseteq 2^X \backslash \{\varnothing\}$ (the budgets), and a *nonempty* choice function $c : \mathcal{B} \to 2^X \backslash \{\varnothing\}$ (satisfying the usual properties), for which all predicates are interpreted properly.

Define the theory of rationalizable choice, T_R, to be the subtheory of T_C where, for each structure $\mathcal{M}$, the associated choice function is rationalizable by a weak order.[7] That is, there exists a weak order R on the global set of alternatives X for which $c(B) = \{x \in B : \forall y \in B, x\, R\, y\}$.

The following theorem is well-known (for example, see Richter (1966)), but we establish it here using our framework.

11. **Proposition.** *T_R is universally and r.e. axiomatizable with respect to T_C.*

Proof. Introduce the language $\mathcal{L} = \langle A, B, \in, C, R\rangle$, where all predicates $A, B, \in, C$ are as in $\mathcal{F}$, and R is a binary predicate. Consider the $\mathcal{L}$-theory T axiomatized by the sentences:

(1) $\forall x \forall y \forall z (\in(x,z) \wedge \in(y,z) \wedge C(x,z)) \to R(x,y)$
(2) $\forall x \forall y \forall z (\in(x,z) \wedge \in(y,z) \wedge R(x,y) \wedge C(x,z)) \to C(x,z)$
(3) $\forall x \forall y (A(x) \wedge A(y)) \to (R(x,y) \vee R(y,x))$
(4) $\forall x \forall y \forall z (A(x) \wedge A(y) \wedge A(z)) \to ((R(x,y) \wedge R(y,z) \to R(x,z)))$

Note that any structure $\mathcal{M} \in T_C$ is a member of T_R if and only if there exists a binary relation R on the global set of alternatives for which for all budgets $B \in \mathcal{B}$, $x \in c(B) \to x\, R\, y \forall y \in B$ and $\forall x \in B$, if $y \in c(B)$ and $x\, R\, y$, then $x \in c(B)$. To see this, note that if R rationalizes c, then clearly the preceding two conditions are satisfied for R. On the other hand, suppose these conditions are satisfied for some R. We claim that R rationalizes c. To see this, note that if $B \in \mathcal{B}$, $x, y \in B$, and $x \in c(B)$, then clearly $x\, R\, y$. On the other hand, suppose that $x\, R\, y$ for all $y \in B$. Then because c is nonempty, there exists $y^* \in c(B)$. Then in particular $x\, R\, y^*$. Conclude $x \in c(B)$.

[7] A weak order is complete and transitive.

Now note that $T_R = F(T) \cap T_C$. And as T is universally axiomatizable, we conclude that T_R is universally and r.e. axiomatizable with respect to T_C by Theorem 5 and Corollary 8. $\square$

3. Rationalizing group preferences

In Section 2.4 we used classical revealed preference theory to illustrate how our theorem can be applied. The results were already known, each having been obtained with arguments designed for each special case. We now turn to apply our theorem to obtain new results for more recent theories of choice. We present an application of our result to theories of group preferences. This section can be understood as an exercise in characterizing the empirical content of collective decision making.

The simplest model looks at the Pareto relation with two agents. Consider the language $\mathcal{F} = \langle R \rangle$ with one binary predicate; $R(x, y)$ is intended to mean that x is preferred to y.

We we are interested in the theory of all structures (X, R^X), where R^X is a partial order, for which there exists *two* complete and transitive binary relations $\succeq_1$ and $\succeq_2$ satisfying the axiom:

$$\forall x \forall y (R(x, y) \leftrightarrow (\succeq_1 (x, y) \wedge \succeq_2 (x, y))). \tag{1}$$

The axiom states that R is the Pareto relation associated to the pair of preferences $\succeq_1$ and $\succeq_2$. We can term this theory the theory of Pareto relations of groups of size 2.

The language $\mathcal{F}$ might allow us to write the axioms on R that describe the theory. Dushnik and Miller (1941) give one such axiomatization, but it is in general a difficult problem to describe this theory using only formulas from R. Our results imply the existence of a recursively enumerable universal axiomatization. Consider the language $\mathcal{L} = \{R, \succeq_1, \succeq_2\}$ and the $\mathcal{L}$-theory axiomatized by axiom (1), together with axioms for $\succeq_i$ being complete and transitive. Theorem 5 and Corollary 8 imply that the theory of Pareto relations of groups of size 2 has a universal and r.e $\mathcal{F}$-axiomatization.

Our results are applicable beyond the theory of Pareto relations. We turn to problems of group preference, where one may hypothesize that group has a single preference which is derived from the application of some specific preference aggregation rule. The idea is that each agent in a group has a "rational" preference, and that some aggregation procedure determines the preferences of the group. Given data on the group's choices, we want to test whether the hypothesis of preference aggregation is true.

In this model, given is a fixed and finite set of agents, each of whose preferences are unobservable, but hypothesized to be rational. Given is also a rule for aggregating agents' preferences into a single preference. The aggregate preference is simply a ranking specified by the social choice rule. @e observe an aggregate preference (a revealed preference), and we would like to know whether it could be generated by the rule for *some* profile of agents' preferences.[8] We want to test whether or not a specific group of agents uses a particular preference aggregation rule in making decisions, only having observed the aggregate ranking. This question is the correct formulation of the standard revealed-preference exercise for the group preference model.

Group preference and social choice theories are an excellent example of how hard it can be to show falsifiability. The theories have a trivial existential (second-order) axiomatization: Given a preference aggregation rule, the theory is the collection of observables for which *there exist* preferences for individuals generating the observable behavior. This second-order axiomatization is the "as if" RP formulation of the theory that we referred to in the introduction. It does not provide an effective test of the theory any more than the analogous formulation of individual utility maximization.

Our main assumptions are that the cardinality of the agents is fixed and finite, and that our preference aggregation rule is neutral and satisfies independence of irrelevant alternatives. We show that any such theory has a universal axiomatization.

[8]In this paper, we focus on preferences which are linear orders; however the results apply more broadly.

The previous literature relates to the theory of social choice, where, given some preference aggregation rule, there have been efforts to axiomatize relations which are rationalized by some society of agents. When the society can be arbitrarily large, it is known that *any* transitive antisymmetric relation is the Pareto relation for *some* society (which may be infinite). This observation is due to Szpilrajn (1930) and Dushnik and Miller (1941)–see also Donaldson and Weymark (1998) and Duggan (1999) for related results in the economics literature. Because of this, any complete binary relation with a transitive asymmetric part is the result of the Pareto extension rule for some society (we identify indifferent alternatives for the Pareto extension rule with unranked alternatives for the Pareto ordering–see Sen (1969)). By contrast, results for the Pareto extension rule with a fixed and finite set of agents are much more scarce and are generally only understood for the two-agent case. Dushnik and Miller (1941) give necessary and sufficient conditions for a binary relation to be the intersection of a pair of linear orders, but this axiomatization is second-order existential. Universal first order axiomatizations based on the Dushnik and Miller (1941) axiomatization are known, and aside from basic completeness and transitivity requirements, boil down to to the absence of certain types of cycles of the indifference relation (see in particular Section 3.2 of Trotter (2002) and the references therein–also Baker, Fishburn, and Roberts (1972)). In the economics literature, Sprumont (2001) provides a similar characterization in a case where preferences are restricted to be suitably "economic." Knoblauch (2005) establishes that there is no axiomatization which relies on a bounded (finite) number of variables. Finally, Echenique and Ivanov (2011) provide an existential characterization in an abstract choice environment.

Results for majority rule are even weaker: McGarvey (1953) showed that any complete binary relation is the majority rule relation for some society of agents (which again may be large). Deb (1976) extends the result to more social choice rules; Kalai (2004) generalizes this result to an even broader class. Shelah (2009) establishes results on different domains of preferences of individuals. But we know of no results characterizing such relations for a fixed set of agents.

We work with neutral preference aggregation rules which satisfy independence of irrelevant alternatives. By working with such preference aggregation rules, we need not specify what the global set of alternatives is in advance. A set of agents N is fixed and finite. A *preference aggregation rule* is therefore defined to be a mapping carrying any set of alternatives X and any N vector of linear orders[9] (termed a *preference profile*) over those alternatives $(\succeq_1, ..., \succeq_n)$ to a complete binary relation over X. We write $R_{f(\succeq_1,...,\succeq_n)}$ for the binary relation which results (suppressing notation for dependence on X). We assume the following property:

12. **Definition.** (Neutrality and Independence of irrelevant alternatives): For all sets X and Y, for all $x, y \in X$ and all $w, z \in Y$ and all preference profiles $(\succeq_1, ..., \succeq_n)$ over X and $(\succeq'_1, ..., \succeq'_n)$ over Y, if for all $i \in N$, $x \succeq_i y \Leftrightarrow w \succeq'_i z$, then $x \, R_{f(\succeq_1,...,\succeq_n)} \, y \Leftrightarrow w \, R_{f(\succeq'_1,...,\succeq'_n)} \, z$.

Our assumption embeds the standard hypotheses of neutrality and independence of irrelevant alternatives. Neutrality means that social rankings should be independent of the names of alternatives, while independence of irrelevant alternatives means that the social preference between a pair of alternatives should depend only on the individual preferences between that pair.

Given f, we will say that a binary relation R on a set X is f-*rationalizable* if there exists a profile of linear orders $(\succeq_1, ..., \succeq_n)$ for which $R = R_{f(\succeq_1,...,\succeq_n)}$.

Let $\mathcal{F} = \langle R \rangle$ be a language involving one binary relation symbol. Given f, a structure (X, R^X) is f-*rationalizable* if R^X is f-rationalizable. Call the class of such $\mathcal{F}$-structures the theory of f-rationalizable preference, or T_f.

13. **Theorem.** *For any f satisfying neutrality and independence of irrelevant alternatives, T_f is universally and r.e. axiomatizable.*

[9] A linear order is complete, transitive, and anti-symmetric

Proof. Since f is neutral and satisfies independence of irrelevant alternatives, and outputs complete rankings, we can conclude that there is a collection of sets $\mathcal{N}_f \subseteq 2^N$, satisfying $E \notin \mathcal{N}$ implies $N \backslash E \in \mathcal{N}_f$, for which for any set X, any profile of linear orders $(\succeq_1, ..., \succeq_n)$ over X, and any pair $x, y \in X$, $x \, R_{f(\succeq_1,...,\succeq_n)} \, y$ if and only if either $x = y$, or there is $E \in \mathcal{N}_f$ for which for all $i \in E$, $x \succeq_i y$, and for all $i \notin E$, $\neg(x \succeq_i y)$.[10]

Consider the language $\mathcal{L} = \langle R, \succeq_1, ..., \succeq_n \rangle$, and the $\mathcal{L}$-theory T axiomatized by the following sentences:

Completeness of $\succeq_i$:

$$\forall x \forall y \succeq_i (x, y) \vee \succeq_i (y, x).$$

Transitivity of $\succeq_i$:

$$\forall x \forall y \forall z \succeq_i (x, y) \wedge \succeq_i (y, z) \rightarrow \succeq_i (x, z).$$

Antisymmetry of $\succeq_i$:

$$\forall x \forall y \succeq_i (x, y) \wedge \succeq_i (y, x) \rightarrow x = y.$$

Finally, f-rationalizability:

$$\forall x \forall y R(x, y) \leftrightarrow (x = y) \vee \bigvee_{E \in \mathcal{N}_f} \left(\bigwedge_{i \in E} (\succeq_i (x, y)) \wedge \bigwedge_{i \notin E} (\neg \succeq_i (x, y)) \right).$$

Finally, note that $T_f = F(T)$, so the result follows by Theorem 5 and Corollary 8.

□

14. *Remark.* The above discussion assumes that preferences are linear orders, but many social choice papers put different restrictions on individual preferences. It is easy to see that the theorem is true on different

[10] To see why, simply let X be any binary set $\{x, y\}$. Define $\succeq_E$ to be the preference profile for which for all $i \in E$, $x \succeq_i y$, and for all $i \notin E$, $\neg(x \succeq_i y)$. Define $\mathcal{N}_f = \{E \subseteq 2^N : x \, R_{f(\succeq_E)} \, y\}$. Note that by completeness of $f(\succeq_E)$, if $E \notin \mathcal{N}_f$, then $N \backslash E \in \mathcal{N}_f$. Finally, by neutrality and independence of irrelevant alternatives, the characterization holds across preference profiles and sets X.

domains of preference profiles: any domain in which individual preferences are universally axiomatizable (weak orders, for example) will work. The condition of f-rationalizability in the proof in general may have to be slightly changed to accommodate individual indifference.

4. Rationalizing group choice behavior

In this section, we look at Nash equilibrium behavior. We assume that we observe a collection of game forms, and a choice made from each game form. We ask whether or not there could exist *strict* preferences for a collection of agents over those game forms which generate the observed choices as Nash equilibrium behavior. We show, using Theorem 5, that this theory has a universal axiomatization.

We first have to set up our framework. Instead of focusing on Nash equilibrium specifically, we work with a general collection of theories of group choice. Nash equilibrium, strong Nash equilibrium, and Pareto optimal choice are special cases. We fix a finite set of agents $N = \{1, ..., n\}$ and a collection $\Gamma \subseteq 2^N \backslash \{\varnothing\}$. The elements of Γ are the sets of agents that can deviate from a profile of strategies.

A *game form* is a tuple $(S_1, ..., S_n)$, of nonempty sets, where we think of S_i as the set of strategies available to agent i. For each profile of preferences $(\succeq_1, \dots, \succeq_n)$ over $\prod_{i \in N} S_i$, a game form $(S_1, ..., S_n)$ defines a normal-form game

$$(S_1, ..., S_n, \succeq_1, \dots, \succeq_n).$$

We define a Γ-Nash equilibrium of a game $(S_1, ..., S_n, \succeq_1, ..., \succeq_n)$ to be $s \in \prod_{i \in N} S_i$ for which for all $\gamma \in \Gamma$ and all $s' \in \prod_{i \in N} S_i$, if there exists $j \in \gamma$ for which $(s'_\gamma, s_{-\gamma}) \succ_j s$, then there exists $k \in \gamma$ for which $s \succ_k (s'_\gamma, s_{-\gamma})$. If we think of Γ as a collection of 'blocking' coalitions, a Γ-Nash equilibrium s is a strategy profile whereby no group $\gamma \in \Gamma$ is willing to jointly deviate, where at least one agent $i \in \gamma$ strictly wants to deviate.

The following are special cases:

- Nash equilibrium results when $\Gamma = \{\{i\} : i \in N\}$
- Pareto optimality results when $\Gamma = \{N\}$

- Strong Nash equilibrium results when $\Gamma = 2^N \backslash \{\varnothing\}$

Other kinds of theories are permissible. For example, by setting $\Gamma = \{G : |G| > |N|/2\}$, we get a kind of majority rule core.

We imagine that we observe a collection of game forms, and *some* strategy profiles which are chosen from each. We do not necessarily observe the entire collection of strategy profiles which could potentially be chosen.

We ask when the strategy profiles are rationalizable by a list of preference relations; obviously, if we make no restriction on preferences, then *every* strategy profile is rationalizable by complete indifference. To this end, we require that preferences be *strict* over strategy profiles. This is a significant assumption.

Let us define the *language of group choice* $\mathcal{F}$ to include the following predicates:

- For each $i \in N$, one unary predicate S_i, where $S_i(y)$ is intended to mean that y is a set of strategies for i
- For each $i \in N$, one unary predicate s_i, where $s_i(x)$ means that x is a strategy for i
- The typical set theoretic binary predicate $\in$, meant to signify membership in a set
- A $2n$-ary predicate R, where $R(y_1, ..., y_n, x_1, ..., x_n)$ means that $(x_1, ..., x_n)$ is observed as being chosen from game form $(y_1, ..., y_n)$

The *theory of group choice* T_G is the class of all structures for the preceding language constructed in the following way. For each agent $i \in N$, there is a *global strategy space* $\mathcal{S}_i \neq \varnothing$ for which the following objects are the elements of the universe:

- Each nonempty $S_i \subseteq \mathcal{S}_i$
- Each $s_i \in \mathcal{S}_i$.

The predicates S_i, s_i, $\in$ are all interpreted properly. Finally, for each game form $\prod_i S_i^*$ and strategy profile $(s_1^*, ..., s_n^*)$, $R(S_1^*, ..., S_n^*, s_1^*, ..., s_n^*)$ implies that $S_i(S_i^*)$, $s_i(s_i^*)$, and lastly that $s_i^* \in S_i^*$. This latter requirement means that only strategy sets can go in the first n

places in R, and that only strategies can go in the last n places. $R(S_1^*, ..., S_n^*, s_1^*, ..., s_n^*)$ means that strategy profile $(s_1^*, ..., s_n^*)$ is chosen from game form $(S_1^*, ..., S_n^*)$–this explains the requirement that $s_i^* \in S_i^*$.

The *theory of* Γ*-rationalizable choice* $T_\Gamma \subseteq T_G$ is the theory of group choice for which for each $i \in N$, there exists a linear order $\succeq_i$ over $\prod_{i\in N} \mathcal{S}_i$ for which for all game forms $(S_1^*, ..., S_n^*)$, $R(S_1^*, ..., S_n^*, s_1^*, ..., s_n^*)$ implies that $(s_1^*, ..., s_n^*)$ is a Γ-Nash equilibrium of the normal-form game $(S_1^*, ..., S_n^*, \succeq_1, \dots, \succeq_n)$.

15. **Theorem.** *The theory of* Γ*-rationalizable choice is universally and recursively enumerably axiomatizable with respect to the theory of group choice.*

Note that Theorem 15 deals with the universal axiomatization of Γ-rationalizable choice with respect to the theory of group choice. We do not here want to focus on axiomatizing group choice; we want to focus only on the additional empirical content of Γ-Nash equilibrium.

Proof. Consider the language $\mathcal{L}$ which includes all predicates in $\mathcal{F}$, but also includes, for each agent i, a $2n$-ary predicate $\succeq_i$.

We use the abbreviation

$$\succ_k (x_1, ..., x_n, z_1, ..., z_n) = \succeq_k (x_1, ..., x_n, z_1, ..., z_n) \wedge \neg \left(\bigwedge_{i=1}^{n} x_i = z_i \right).$$

For each $\gamma \in \Gamma$ and $k \in \gamma$, we use the following shorthand:

If $|\gamma| > 1$,

$$\varphi_{\gamma,k}(x_1, ..., x_n, z_1, ..., z_n) =$$
$$\succ_k ((z_\gamma, x_{-\gamma}), x) \to \bigvee_{i\in\gamma\setminus\{k\}} \succ_i (x, (z_\gamma, x_{-\gamma}))$$

otherwise, if $|\gamma| = 1$,

$$\varphi_{\gamma,k}(x_1, ..., x_n, z_1, ..., z_n) = \succeq_k (x_1, ..., x_n, z_1, ..., z_n).$$

Consider the theory T axiomatized by the following sentences.

For each $\gamma \in \Gamma$ and $k \in \gamma$,

$$\forall x_1 ... \forall x_n \forall y_1 ... \forall y_n \forall z_1 ... \forall z_n$$
$$\bigwedge_{i \in \gamma} \in (z_i, y_i) \wedge \bigwedge_{i \in N} \in (x_i, y_i) \wedge R(y_1, ..., y_n, x_1, ..., x_n)$$
$$\to \varphi_{\gamma,k}(x_1, ..., x_n, z_1, ..., z_n)$$

Completeness: For each $k \in N$,

$$\forall s_1 ... \forall s_n \forall s'_1 ... \forall s'_n$$
$$\bigwedge_{i=1}^{n} s_i(s_i) \wedge \bigwedge_{i=1}^{n} s_i(s'_i) \to (\succeq_k (s_1, ..., s_n, s'_1, ..., s'_n) \vee \succeq_k (s'_1, ..., s'_n, s_1, ..., s_n))$$

Transitivity:

$$\forall s_1 ... \forall s_n \forall s'_1 ... \forall s'_n \forall s''_1 ... \forall s''_n$$
$$\bigwedge_{i=1}^{n} s_i(s_i) \wedge \bigwedge_{i=1}^{n} s_i(s'_i) \wedge \bigwedge_{i=1}^{n} s_i(s''_i)$$
$$\to (\succeq_k (s_1, ..., s_n, s'_1, ..., s'_n) \wedge \succeq_k (s'_1, ..., s'_n, s''_1, ..., s''_n) \to \succeq_k (s_1, ..., s_n, s''_1, ..., s''_n))$$

Antisymmetry:

$$\forall s_1 \forall s_2 ... \forall s_n \forall s'_1 ... \forall s'_n$$
$$\left(\bigwedge_{i=1}^{n} s_i(s_i) \wedge \bigwedge_{i=1}^{n} s_i(s'_i) \right) \wedge$$
$$(\succeq_k (s_1, ..., s_n, s'_1, ..., s'_n) \wedge \succeq_k (s'_1, ..., s'_n, s_1, ..., s_n))$$
$$\to \bigwedge_{i=1}^{n} (s_i = s'_i)$$

As T has a universal axiomatization, so does $F(T)$. Since the axiomatization of T is finite, $F(T)$ has a recursively enumerable universal axiom by Corollary 8. And $T_\Gamma = F(T) \cap T_G$. So T_Γ has a r.e. universal axiomatization within T_G.

□

Because we ask for group choice functions to be rationalizable by strict preferences, the exercise here is similar in spirit to Afriat (1967), who assumes partial observations of demand functions and asks demand functions to be rationalizable by locally non-satiated preference.

5. Complexity results

We now turn to a discussion of the complexity of the theory $F(T)$, for a given theory T. In Corollary 8, we addressed the computational structure of the set of axioms of $F(T)$, and concluded that this was a r.e. set under appropriate conditions: this means that there is an algorithm that can determine that a data set falsifies the theory. Here, we instead ask about the computational structure of the theory $F(T)$ itself, as opposed to its axioms. To render this a meaningful exercise, we here restrict to finite structures.

The question addressed is as follows. Given a finite $\mathcal{F}$-structure $\mathcal{M}$, how difficult, or *complex* is it to determine that $\mathcal{M} \in F(T)$? This is therefore the problem of *verifying* whether or not a structure is a model of our theory. In contrast; if we take interest in falsification, we may ask, given a finite $\mathcal{F}$-structure $\mathcal{M}$, how difficult is it to determine whether $\mathcal{M} \notin F(T)$?

We say that a theory $F(T)$ is in class P (for polynomial time) if there is a Turing machine operating in polynomial time that, given any $\mathcal{F}$-structure $\mathcal{M}$, tells us whether $\mathcal{M} \in F(T)$. We say that a theory $F(T)$ is in class NP (for non-deterministic polynomial time), if there is a non-deterministic Turing machine (essentially a multi-tape Turing machine) running in polynomial time that, given any $\mathcal{F}$-structure $\mathcal{M}$, tells us whether $\mathcal{M} \in F(T)$. Equivalently, a theory is in NP if there is a polynomial time (deterministic) Turing machine such that for any $\mathcal{M} \in F(T)$, there is a *certificate* (depending on $\mathcal{M}$) which can be used to verify that $\mathcal{M} \in F(T)$. For this reason, the class NP is often understood as the class of theories for which there is a polynomial time *proof* that a structure is in the theory.[11]

The following theorem is due to Fagin (1974), and can be found in many references on descriptive complexity, for example Immerman (1999), Papadimitriou (2003), or Libkin (2004).

16. **Theorem** (Fagin). *Suppose that T is a finitely axiomatized $\mathcal{L}$ theory. Then $F(T)$ is in NP.*

In fact, Fagin's theorem is much deeper than this (it is an equivalence theorem, of which we have provided the simple part), but some intuition for this direction can be given. Essentially, if we have a finite structure $\mathcal{M}$, to determine that $\mathcal{M} \in F(T)$, we simply need to find $\mathcal{M}' \in T$ for which $\mathcal{M} = F(\mathcal{M}')$. $\mathcal{M}'$ will have associated with it relations for predicates in $\mathcal{L}\backslash\mathcal{F}$: checking that these relations satisfy the finite first order axioms of T provides a polynomial time "proof" that $\mathcal{M} \in F(T)$. The non-determinism of the Turing machine allows it to "guess" the nature of the additional relations in $\mathcal{M}'$.

As a consequence of Fagin's theorem, and the formalism introduced in our paper, we can conclude that the problem of verifying that a finite structure is a model of rational choice, group preference or Nash equilibrium are all in NP.

While we can establish that $F(T)$ is in NP, it is often hard to say more. General results on these issues do exist, and it is known that certain "types" of axiomatizations lead to different complexity classes. Excellent references include Immerman (1999) and Libkin (2004). We will simply say that there are examples of theories T which generate $F(T)$ which is in P and theories T which generate $F(T)$ which is NP-complete.[12]

[11]Dually, the class of theories for which there is a polynomial time algorithm which, given a structure not in a theory, determines it is not in the theory is called co-NP. The class co-NP is clearly more relevant for falsification purposes.

[12]These are the "hardest" problems in NP to solve.

6. Large sample results

A related issue dealing with finite structures is the *proportion* of structures which are in $F(T)$. Suppose we fix a cardinality n. We can ask about the proportion of $\mathcal{F}$-structures with universe $\{1, ..., n\}$ which are elements of $F(T)$.[13] Denote this number by $\mu^{F(T)}(n)$. The following simple result can be established.

17. **Theorem.** *Suppose that T is a universal $\mathcal{L}$ theory, and that there exists $\mathcal{M} \notin F(T)$. Then $\mu^{F(T)}(n) \to 0$.*

Proof. Suppose that T is a universal $\mathcal{L}$ theory, and that there exists $\mathcal{M} \notin F(T)$. Then by Theorem 5, $F(T)$ is a universal $\mathcal{F}$ theory. Since $\mathcal{M} \notin F(T)$, by theorem 1.2 of Tarski (1954), there is a finite structure $\mathcal{M}' \notin F(T)$. Suppose this structure has cardinality m; without loss of generality, we may assume it has universe $\{1, ..., m\}$ (since universal theories are closed under isomorphism).

Since $\mathcal{M}' \notin F(T)$, there is a universal axiom σ of $F(T)$ which is not satisfied by $\mathcal{M}'$. Let the $\mathcal{F}$ theory axiomatized by σ be denoted T^σ. Let $\mu^{T^\sigma}(n)$ be the proportion of structures on $\{1, ..., n\}$ which satisfy σ; we have established that $\mu^{T^\sigma}(m) = p < 1$.

Let a be an integer, and consider the set of structures with universe $\{1, ..., am\}$ which satisfy σ. For any $k \in \{0, a-1\}$, the probability that a structure satisfies σ when the universal quantifiers are restricted only to $\{k+1, ..., k+m\}$ is p; consequently, $\mu^{T^\sigma}(am) \leq p^a$.

Finally, consider any l, and consider the set of all structures on $\{1, ..., l\}$. Let $a(l)$ be such that $a(l)m \leq l$. We can again ask which proportion of structures on $\{1, ..., l\}$ satisfy the axiom σ when its quantifiers are restricted only to $\{1, ..., a(l)m\}$. Clearly, this number is equal to $\mu^{T^\sigma}(a(l)m)$. And requiring σ to hold for all elements of $\{1, ..., l\}$ therefore results in a smaller proportion of structures: consequently, $\mu^{T^\sigma}(l) \leq \mu^{T\sigma}(a(l)m) \leq p^{a(l)}$. As $a(l)$ increases with l, we conclude that $\mu^{T^\sigma}(l) \to 0$, consequently $\mu^{F(T)}(l) \to 0$. $\square$

[13] Since our theories are closed under isomorphisms, what matters is that the universe of the structure has cardinality n.

The result roughly states that any universally quantified revealed preference theory will tend to have very few models for large universes. The result is very general, but is perhaps unsurprising when thought of dually. A universal quantifier is the negation of an existential quantifier: therefore, universal quantification rules out existence of certain objects. As we have larger and larger models, the likelihood that at least one object exists which has been ruled out necessarily increases. For example, the axiom of acyclicity (ruling out cycles of finite length for binary relations) becomes harder and harder to satisfy: with large universes, there are more and more possible cycles.

Some care is needed in interpreting the result, as it says nothing of the probability of *actually* observing structures which are or are not elements of $F(T)$. To discuss such issues, one would need to talk about the process which generates structures. This interpretation is only meaningful if we believe that structures are drawn according to a uniform distribution.

7. Discussion and relation to previous literature

7.1. **Philosophy of science.** The type of issues we discuss here have previously been studied by philosophers of science. Without going into full detail, Ramsey (1931) was one of the first to discuss the elimination of "theoretical" terms from scientific theories. Various authors give different interpretation to the notion of "Ramsey elimination." The work of Sneed (1971) includes notions very similar to ours; in particular, he defines a finitely axiomatized $\mathcal{L}$- theory T to be Ramsey eliminable if $F(T)$ is a finitely axiomatized $\mathcal{F}$-theory. In particular, he includes an example (attributed to Dana Scott) of a theory T which is first order axiomatizable, but for which $F(T)$ is not first order axiomatizable. We include here an adaptation of this example.

18. **Example.** Let $\mathcal{F} = \langle R \rangle$, where R is a unary predicate, and let $\mathcal{L} = \langle R, Q \rangle$, where Q is a binary predicate. Consider the $\mathcal{L}$-theory T axiomatized by the following sentences:

(1) $\forall x \forall y Q(x, y) \to R(x) \wedge \neg R(y)$

(2) $\forall x R(x) \rightarrow (\exists y Q(x,y) \wedge (\forall z Q(x,z) \rightarrow y = z))$
(3) $\forall x \neg R(x) \rightarrow (\exists y Q(y,x) \wedge (\forall z Q(z,x) \rightarrow y = z))$

T is the theory of all structures (X, R^X, Q^X) for which there is a one-to-one correspondence Q^X between the elements of R^X and its complement.

19. **Proposition.** *$F(T)$ is not first order $\mathcal{F}$-axiomatizable.*

Proof. Suppose by means of contradiction that there is a first order axiomatization of $F(T)$. Consider any $\mathcal{F}$ structure (X, R^X) where $|R^X|$ is infinite, $|X \backslash R^X|$ is infinite, and the cardinalities of R^X and $X \backslash R^X$ are distinct. Note that $(X, R^X) \notin F(T)$.

By the Löwenheim-Skolem theorem, (see for example Marker (2002), Theorem 2.3.7), there exists a countable structure $(X', R^{X'})$ which satisfies exactly the same first order sentences as (X, R^X). But note in particular that for each $n > 0$, the sentence

$$\exists x_1 ... \exists x_n \bigwedge_{i=1}^{n} R(x_i) \wedge \bigwedge_{i \neq j} (\neg(x_i = x_j))$$

is valid for (X, R^X); in particular, then, it is valid for $(X', R^{X'})$; consequently, $|R^{X'}|$ is infinite. Similarly, since

$$\exists x_1 ... \exists x_n \bigwedge_{i=1}^{n} \neg R(x_i) \wedge \bigwedge_{i \neq j} (\neg(x_i = x_j))$$

is valid for (X, R^X), it is also valid for $(X', R^{X'})$ and hence $|X' \backslash R^{X'}|$ is infinite. Since X' is countable, there is therefore a bijection between $R^{X'}$ and $X \backslash R^{X'}$, so that $(X', R^{X'}) \in F(T)$. But then (X, R^X) satisfies the sentences axiomatizing $F(T)$ (as it satisfies the same sentences as $(X', R^{X'})$ and $(X', R^{X'}) \in F(T)$). So $(X, R^X) \in F(T)$, a contradiction. $\square$

The preceding example is important in that it illustrates that the problem of axiomatizability of $F(T)$ is non-trivial. To this end, Van Benthem (1978) (Theorem 4.2) uncovers necessary and sufficient conditions for $F(T)$ to be first order axiomatizable, given that T is first order axiomatizable. His conditions are essentially an adaptation of a well-known theorem in model theory axiomatizing those theories which are first order axiomatizable (see, for example Chang and Keisler (1990) Theorem 4.1.12). The condition is a semantic condition requiring one to verify, for any model of $F(T)$, whether any structure which satisfies exactly the same $\mathcal{F}$-sentences is also a model of $F(T)$. In practice, this is nearly impossible to verify. Our condition; on the other hand, is a syntactic condition which is trivial to verify.

7.2. **Economics.** The questions studied here belong, generally speaking, in the realm of the philosphy of science, but we believe that the formalism adopted here (and borrowed from model theory in mathematics) is particularly appealing to economic theorists. We have demonstrated this fact by using familiar examples from both classical and recent research in economic theory. Our feeling is that few other disciplines have embraced the axiomatic method as economics has; our results are unlikely to have interesting applications in, say, biology or chemistry.

That formal methodological questions are appealing to economists is reflected in the interest taken by other economists in these questions. Herbert Simon wrote a sequence of papers on falsifiability and empirical content. For example, Simon (1985) discusses some of the issues we discuss here: Simon argues that the theory of rational choice is falsifiable, even though the RP formulation of the theory is existentially quantified over unobservables. As Simon (1985) states, "although existential quantification of an observable is fatal to the falsifiability of a theory, the same is not true when the existentially quantified term is a theoretical one."

While this may seem obvious to many, it has led to a large degree of confusion among economists. For example Boland (1981) argued that the theory of rational choice is not falsifiable precisely because of its existential formulation over unobservables. In his words:

> Given the premise–"All consumers maximize something"–the critic can claim he has found a consumer who is not maximizing anything. The

> person who assumed the premise is true can respond: "You claim you have found a consumer who is not a maximizer but how do you know there is not something which he is maximizing?" In other words, the verification of the counterexample requires the refutation of a strictly existential statement; and as stated above, we all agree that one cannot refute existential statements.

Mongin (1986) beautifully counters this argument. In this context, he has already observed that all $\mathcal{F}$-implications of T are $\mathcal{F}$-implications of $F(T)$. It follows from this that if T is universal, then $F(T)$ has universal implications, and is hence falsifiable. As noted above, we have gone further than this in showing that in fact, $F(T)$ is first-order axiomatizable (indeed, universally axiomatizable). Hence *all* of its implications are falsifiable.

Our work is related to the approach in Brown and Matzkin (1996), and the general approach to testable implications discussed in Brown and Kubler (2008). In these papers, as in ours, there is an operation of projection to eliminate certain existential quantifications. The idea in Brown-Matzkin and in Brown-Kubler is to exploit the property of quantifier elimination in certain mathematical theories. Our work, on the other hand, uses results from model theory on when a universal axiomatization is possible. Our projection argument follows from the verification that the universal axiomatization can be projected. We do not exploit the property of quantifier elimination of the mathematical theory underlying the economic model (indeed our results may apply when quantifier elimination does not hold).

In our previous paper, Chambers, Echenique, and Shmaya (2010), we dealt with theories which could be axiomatized by what we called *UNCAF* formulas, for universal negation of conjunction of atomic formulas.[14] Under certain conditions, being UNCAF axiomatizable is equivalent to being universally axiomatizable. And in fact, we show there that a result akin to Theorem 5 is true for UNCAF sentences.

[14]The strong axiom of revealed preference described above is an example of a collection of such formulas.

This proof relies on drastically different techniques, as there are deep differences between theories axiomatizable by UNCAF sentences and those which are only universally axiomatizable. In particular, theories which are UNCAF axiomatizable are closed under weak substructures (and not just substructures), a property that plays a critical role in our previous paper.

The reason Theorem 5 is useful is because, in general, the hypothesized theory T is usually not axiomatizable by UNCAF sentences. For example, the axioms of completeness and transitivity for binary relations have no UNCAF axiomatization. In general, then, it is impossible to obtain any results about the falsifiability of a revealed preference theory without having some result about universality.

Appendix A. Basic definitions from Model Theory

The following definitions are taken, for the most part, quite literally from (Marker, 2002), pp. 8-12. We refer readers to this excellent text for more details; but present the basics here to keep the analysis self-contained. The $\overline{x}$ notation is here used to denote a list, or vector, or elements $(x_1, ..., x_m)$.

20. **Definition.** $\mathcal{M}$ is a *substructure* of $\mathcal{N}$ if $M \subseteq N$ and the inclusion map $\iota : M \to N$ defined by $\iota(m) = m$ for all $m \in M$ is an $\mathcal{L}$-embedding.

The following definition gives us the basic building blocks of our syntax. Note that we include a countable list of "variables" to be used in this definition; these are not part of the language *per se*, but rather part of a "meta language" in that they are present in all languages.

21. **Definition.** The set of $\mathcal{L}$-*terms* is the smallest set $\mathcal{TE}$ such that

(1) $c \in \mathcal{TE}$ for each constant symbol $c \in \mathcal{C}$
(2) each variable symbol $v_i \in \mathcal{TE}$ for $i = 1, 2, ...,$
(3) if $t_1, ..., t_{n_f} \in \mathcal{TE}$ and $f \in \mathcal{F}$, then $f(t_1, ..., t_{n_f}) \in \mathcal{TE}$.

22. **Definition.** Say that ϕ is an *atomic $\mathcal{L}$-formula* if ϕ is one of the following

(1) $t_1 = t_2$, where t_1 and t_2 are terms

(2) $R(t_1, ..., t_{n_R})$, where $R \in \mathcal{R}$ and $t_1, ..., t_{n_R}$ are terms

23. **Definition.** The set of *$\mathcal{L}$-formulas* is the smallest set $\mathcal{W}$ containing the atomic formulas such that

(1) if ϕ is in $\mathcal{W}$, then $\neg\phi$ is in $\mathcal{W}$
(2) if ϕ and ψ, then $(\phi \wedge \psi)$ and $(\phi \vee \psi)$ are in $\mathcal{W}$
(3) if ϕ is in $\mathcal{W}$, then $\exists v_i \phi$ and $\forall v_i \phi$ are in $\mathcal{W}$.

24. **Definition.** A variable v *occurs freely* in a formula ϕ if it is not inside a $\exists v$ or $\forall v$ quantifier. It is *bound* in ϕ if it does not occur freely in ϕ.

25. **Definition.** A *sentence* is a formula ϕ with no free variables.

We are now prepared to define a concept of "truth" relating syntax and semantics. We want to define what it means for a sentence to be true in a given structure.

26. **Definition.** Let ϕ be a formula with free variables from $\overline{v} = (v_{i_1}, ..., v_{i_m})$, and let $\overline{a} = (a_{i_1}, ..., a_{i_m}) \in M^m$. We inductively define $M \models \phi(\overline{a})$ as follows. The notation $M \not\models \psi(\overline{a})$ means that $M \models \phi(\overline{a})$ is not true.

(1) If ϕ is $t_1 = t_2$, then $\mathcal{M} \models \phi(\overline{a})$ if $t_1^{\mathcal{M}}(\overline{a}) = t_2^{\mathcal{M}}(\overline{a})$
(2) If ϕ is $R(t_1, ..., t_{n_R})$, then $\mathcal{M} \models \phi(\overline{a})$ if $(t_1^{\mathcal{M}}(\overline{a}), ..., t_{n_R}^{\mathcal{M}}(\overline{a})) \in R^{\mathcal{M}}$
(3) If ϕ is $\neg\psi$, then $\mathcal{M} \models \phi(\overline{a})$ if $\mathcal{M} \not\models \psi(\overline{a})$
(4) If ϕ is $(\psi \wedge \theta)$, then $\mathcal{M} \models \phi(\overline{a})$ if $\mathcal{M} \models \psi(\overline{a})$ and $\mathcal{M} \models \theta(\overline{a})$
(5) If ϕ is $(\psi \vee \theta)$, then $\mathcal{M} \models \phi(\overline{a})$ if $\mathcal{M} \models \psi(\overline{a})$ or $\mathcal{M} \models \theta(\overline{a})$
(6) If ϕ is $\exists v_j \psi(\overline{v}, v_j)$, then $\mathcal{M} \models \phi(\overline{a})$ if there is $b \in M$ such that $\mathcal{M} \models \psi(\overline{a}, b)$
(7) If ϕ is $\forall v_j \psi(\overline{v}, v_j)$, then $\mathcal{M} \models \phi(\overline{a})$ if for all $b \in M$, $\mathcal{M} \models \psi(\overline{a}, b)$.

27. **Definition.** $\mathcal{M}$ *satisfies* $\phi(\overline{a})$ or $\phi(\overline{a})$ *is true* in $\mathcal{M}$ if $\mathcal{M} \models \phi(\overline{a})$.

Lastly, for our purposes, it is useful to have a notion of a *universal* sentence.

28. **Definition.** A *universal sentence* or *universal formula* is a sentence of the form $\forall\overline{v}\phi(\overline{v})$, where ϕ is quantifier free.

References

Afriat, S. N. (1967): "The construction of utility functions from expenditure data," *International Economic Review*, 8(1), 67–77.

Ambrus, A., and K. Rozen (2008): "Revealed conflicting preferences," working paper.

Aumann, R. (1960): "Acceptable points in games of perfect information," *Pacific Journal of Mathematics*, 10(2), 381–417.

Baker, K., P. Fishburn, and F. Roberts (1972): "Partial orders of dimension 2," *Networks*, 2(1), 11–28.

Boland, L. (1981): "On the futility of criticizing the neoclassical maximization hypothesis," *The American Economic Review*, 71(5), 1031–1036.

Brown, D., and F. Kubler (2008): "Refutable Theories of Value," in *Computational Aspects of General Equilibrium Theory: Refutable Theories of Value*. Springer Publishing Company, Incorporated.

Brown, D. J., and R. L. Matzkin (1996): "Testable restrictions on the equilibrium manifold," *Econometrica*, 64(6), 1249–1262.

Chambers, C. P., F. Echenique, and E. Shmaya (2010): "The axiomatic structure of empirical content," working paper.

Chang, C., and H. Keisler (1990): *Model theory*. North Holland.

Craig, W. (1953): "On axiomatizability within a system," *The Journal of Symbolic Logic*, 18(1), 30–32.

Deb, R. (1976): "On constructing generalized voting paradoxes," *The Review of Economic Studies*, 43(2), 347–351.

DeClippel, G., and K. Eliaz (2009): "Reason-Based Choice: A Bargaining Rationale for the Attraction and Compromise Effects," working paper.

Donaldson, D., and J. Weymark (1998): "A quasiordering is the intersection of orderings," *Journal of Economic Theory*, 78(2), 382–387.

Duggan, J. (1999): "A general extension theorem for binary relations," *Journal of Economic Theory*, 86(1), 1–16.

DUSHNIK, B., AND E. MILLER (1941): "Partially ordered sets," *American Journal of Mathematics*, 63(3), 600–610.

ECHENIQUE, F., AND L. IVANOV (2011): "Implications of Pareto Efficiency for two-agent (household) choice," *Journal of Mathematical Economics*.

FAGIN, R. (1974): "Generalized first-order spectra and polynomial-time recognizable sets," *Complexity of Computation*, 7, 43–73.

FUDENBERG, D., AND D. LEVINE (2006): "A dual-self model of impulse control," *American Economic Review*, 96(5), 1449–1476.

GALAMBOS, A. (2009): "The complexity of Nash rationalizability," working paper.

GREEN, J. R., AND D. HOJMAN (2008): "Choice, rationality, and welfare measurement," working paper.

HOUTHAKKER, H. (1950): "Revealed preference and the utility function," *Economica*, pp. 159–174.

IMMERMAN, N. (1999): *Descriptive complexity*. Springer Verlag.

KALAI, G. (2004): "Social indeterminacy," *Econometrica*, 72, 1565–1581.

KNOBLAUCH, V. (2005): "Characterizing Paretian preferences," *Social Choice and Welfare*, 25(1), 179–186.

LEDYARD, J. O. (1986): "The scope of the hypothesis of Bayesian equilibrium," *Journal of Economic Theory*, 39(1), 59 – 82.

LEE, S. (2009): "The testable implications of zero-sum games," Mimeo, California Institute of Technology.

LIBKIN, L. (2004): *Elements of finite model theory*. Springer Verlag.

MANZINI, P., AND M. MARIOTTI (2007): "Sequentially rationalizable choice," *American Economic Review*, 97(5), 1824–1839.

MARKER, D. (2002): *Model theory: an introduction*. Springer Verlag.

MCGARVEY, D. C. (1953): "A theorem on the construction of voting paradoxes," *Econometrica*, 21(4), 608–610.

MONGIN, P. (1986): "Are all-and-some statements falsifiable after all?," *Economics and Philosophy*, 2, 185–195.

O'DONOGHUE, T., AND M. RABIN (1999): "Doing it now or later," *The American Economic Review*, 89(1), 103–124.

Papadimitriou, C. H. (2003): *Computational complexity*. John Wiley and Sons Ltd.

Peleg, B., and S. Tijs (1996): "The consistency principle for games in strategic form," *International Journal of Game Theory*, 25(1), 13–34.

Plott, C. R. (1974): "On Game Solutions and Revealed Preference Theory," Discussion Paper 35, California Institute of Technology.

Popper, K. R. (1959): *The Logic of Scientific Discovery*. Hutchinson.

Ramsey, F. (1931): "Theories," *The Foundations of Mathematics*, pp. 212–236.

Richter, M. K. (1966): "Revealed preference theory," *Econometrica*, 34(3), 635–645.

Rogers, H. (1987): *Theory of Recursive Functions and Effective Computability*. MIT Press, Cambridge, MA.

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

Sen, A. (1969): "Quasi-transitivity, rational choice and collective decisions," *The Review of Economic Studies*, 36(3), 381–393.

Shelah, S. (2009): "What majority decisions are possible," *Discrete Mathematics*, 309(8), 2349–2364.

Simon, H. (1985): "Quantification of theoretical terms and the falsifiability of theories," *The British Journal for the Philosophy of Science*, 36(3), 291.

Sneed, J. (1971): *The Logical Structure of Mathematical Physics*. D. Reidel Publishing Company.

Sprumont, Y. (2000): "On the testable implications of collective choice theories," *Journal of Economic Theory*, 93(2), 205–232.

——— (2001): "Paretian quasi-orders: the regular two-agent case," *Journal of Economic Theory*, 101(2), 437–456.

Suzumura, K. (1976): "Remarks on the theory of collective choice," *Economica*, 43(172), 381–390.

Szpilrajn, E. (1930): "Sur l'extension de l'ordre partiel," *Fundamenta Mathematicae*, 16, 386–389.

Tarski, A. (1954): "Contributions to the theory of models I," *Indagationes Mathematicae*, 16, 572–581.

Trotter, W. (2002): *Combinatorics and partially ordered sets: Dimension theory*. Johns Hopkins Univ Pr.

Van Benthem, J. (1978): "Ramsey eliminability," *Studia Logica*, 37(4), 321–336.

Wilson, R. B. (1970): "The finer structure of revealed preference," *Journal of Economic Theory*, 2(4), 348 – 353.

Xu, Y., and L. Zhou (2007): "Rationalizability of choice functions by game trees," *Journal of Economic Theory*, 134(1), 548–556.