To Infinity and Beyond:
A General Framework for Scaling Economic Theories*

Yannai A. Gonczarowski† Scott Duke Kominers‡
Ran I. Shorrer§

May 2, 2024

Abstract

Many economic models incorporate finiteness assumptions that, while introduced for simplicity, play a real role in the analysis. We provide a principled framework for scaling results from such models by removing these finiteness assumptions. Our sufficient conditions are on the theorem statement only, and not on its proof. This results in short proofs, and even allows us to use the same argument to scale similar theorems that were

*An earlier version of this paper, entitled "To Infinity and Beyond: Scaling Economic Theories via Logical Compactness," appeared as a one-page abstract in the *Proceedings of the 21st ACM Conference on Economics and Computation.* We thank David Ahn, Bob Anderson, Morgane Austern, Archishman Chakrabortyz, Chris Chambers, Yunseo Choi, Henry Cohn, Piotr Dworczak, Andrew Ellis, Tamás Fleiner, Drew Fudenberg, Wayne Gao, Jerry Green, Joseph Halpern, Ron Holzman, Ravi Jagadeesan, M. Ali Khan, David Laibson, Rida Laraki, Bar Light, Elliot Lipnowski, Ce Liu, George Mailath, Michael Mandler, Paul Milgrom, Ankur Moitra, Yoram Moses, Juan Pereyra, Marek Pycia, Debraj Ray, John Rehbeck, Phil Reny, Joseph Root, Ariel Rubinstein, Dov Samet, Chris Shannon, Tomasz Strzalecki, Sergiy Verstyuk, Rakesh Vohra, Shing-Tung Yau, Bill Zame, and numerous seminar audiences for helpful comments. Gonczarowski was supported in part by the Adams Fellowship Program of the Israel Academy of Sciences and Humanities; his work was supported in part by ISF grants 1435/14, 317/17, and 1841/14 administered by the Israeli Academy of Sciences; by the United States–Israel Binational Science Foundation (BSF grant 2014389); and by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant No. 740282), and under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC grant number 337122. Kominers gratefully acknowledges the support of the National Science Foundation (grant SES-1459912), as well as the Ng Fund and the Mathematics in Economics Research Fund of the Harvard Center of Mathematical Sciences and Applications. Shorrer was supported by a grant from the United States–Israel Binational Science Foundation (BSF grants 2016015 and 2022417). Part of this work was conducted during the Simons Laufer Mathematical Sciences Institute Fall 2023 program on the Mathematics and Computer Science of Market and Mechanism Design, which was supported by the National Science Foundation under Grant No. DMS-1928930 and by the Alfred P. Sloan Foundation under grant G-2021-16778.

†Department of Economics and Department of Computer Science, Harvard University — *E-mail*: yannai@gonch.name. Parts of the work of Gonczarowski were carried out while at the Hebrew University of Jerusalem, at Tel Aviv University, and at Microsoft Research.

‡Entrepreneurial Management Unit, Harvard Business School; Department of Economics and CMSA, Harvard University; and a16z crypto — *E-mail*: kominers@fas.harvard.edu.

§Department of Economics, Penn State University — *E-mail*: shorrer@psu.edu.

proven using distinctly different tools. We demonstrate the versatility of our approach via an array of examples from revealed-preference theory.

"More is good... all is better."
—Ferengi Rule of Acquisition #242, Star Trek

1 Introduction

In economic theory, we frequently make finiteness assumptions for simplicity and/or tractability—and those assumptions can play a real role in the analysis. Of course, the real world is itself finite, so there is in some sense no "loss" from assuming finiteness in our models. But finiteness assumptions nevertheless sometimes lead to conceptual problems—if our understanding of economic theory hinges on finiteness, then our models may not quite tell the whole story.

For example, in decision theory, revealed preference analysis seeks to understand what we can infer about agents from their choice behavior. While a list of observed choices is always finite, if we make parametric assumptions such as homotheticity, then each data point becomes infinitely many data points. Even without such assumptions, we would like to reason about possible demand functions—defined everywhere—that are consistent with the data, and this requires conjecturing about behavior over an infinite dataset. Furthermore, theorizing about observing infinite datasets lets us separate the limitations of inference about agents' preferences that are just imposed by data finiteness from those that are inherent even with access to every possible observation. As another example, if a game-theoretic finding is true only when the set of agents is finite, then there is an implicit discontinuity, possibly relying on an edge effect or a specific starting condition that may not be robust to small frictions or perturbations.[1] Thus finite-market results that also hold in infinite markets are in some sense more robust.

In this paper, we present a general framework for strengthening results that assume finiteness, by scaling them to infinite settings.[2] Our approach, by relying on results in Propositional Logic, implicitly leverages topological properties of the space of *theorem statements* rather than any features or techniques from their proofs. As such, it allows us to prove results along the lines of "if a certain statement holds when assuming finiteness (regardless of how one would prove it), then—due merely to the structure of this statement—it must hold even if the finiteness assumption is dropped."

Our methods allow us to relax various finiteness assumptions, such as dataset size and market size. In this paper we focus on applications to decision theory, where the infinity that we tackle is the infinity of data. To demonstrate the versatility of our approach across disparate fields, we also demonstrate an application

[1]For an example of a different kind of discontinuity—between a finite and a continuum setting—see the work of Miralles and Pycia (2015), showing that a continuum model may rule out important phenomena that are observed in the finite models that converge to it.

[2]As a side note: economic theory sometimes also turns to infinite models when their finite analogues are hard to analyze—for example, to smooth out integer effects. That is not our focus here.

to game theory, where the infinity that we tackle is that of the market size.

In Section 2, we state and prove our general *scaling lemma* that we apply throughout the paper. Section 3 presents a "warm up," applying our approach to scale a fundamental result for which the proof of the finite case is considerably simpler than that of the infinite case. Specifically, this is that any dataset satisfying the strong axiom of revealed preferences (SARP) is rationalizable, for which the finite case is simple and the infinite case is usually proven by appealing to Zorn's lemma. In Section 4, we proceed with using a proof very similar to the simple proof from the warm-up, to relax a finiteness assumption in a result for which the infinite case has not been previously proven. Specifically, we prove a novel infinite-data version of Masatlioglu et al.'s (2012) characterization of limited-attention rationalizability. The finite-case *proofs* for this result and for the warm-up are starkly different. Nonetheless, the *statements* of these two results are similar, and this enables us to scale them using essentially the same proof. We furthermore show that since our approach only relies on theorem statements (and not on how they are proven) one can use it to conditionally scale a rich family of not-yet-proven results (i.e., conditional on the finite case being true).

At first glance, our approach might seem limited to proving results that are discrete in nature (see discussion below). Nonetheless, in Section 5 we use our approach to prove results regarding objects that are nondiscrete (coming from a continuum space). Specifically, here we reprove Reny's (2015) infinite-data version of Afriat's (1967) theorem, where utilities come from a continuum space, as well as Caplin, Dean, and Leahy's (2017) infinite-data version of Caplin and Dean's (2015) characterization of having a costly information acquisition representation, where priors come from a continuum space. Again, our scaling proofs for both of these theorems are nearly identical. In Section 6, we discuss limitations of our approach in the context of decision theory.

In Section 7, we conclude with an application to a different field and with a different notion of infinity, reproving the existence of a Nash equilibrium in infinite games on graphs. We then discuss limitations of our approach more broadly. Additional results can be found in the working paper (Gonczarowski, Kominers, and Shorrer, 2023).

As already mentioned, some of the results that we prove in this paper are novel to this work. Other results that we (re-)prove have already been obtained using other, very different methods, which allows us to compare and contrast the previous proof techniques with ours. As our illustrative applications demonstrate, proofs that use our framework have several notable features. First, they use **one tool** rather than having to choose from various setting-specific tools. Second, the proof structure is **modular**: our conditions for scaling the finite-case result to the infinite case depend only on the statement of the finite-case result and are completely **agnostic** to the argument/methods used to prove that result. Furthermore, the proofs are **robust** in that even their dependence on the details of the model is quite weak, and essentially the same proof can sometimes be used in quite different models.

1.1 Technique

Our general approach, formulated by Lemma 2.5 in Section 2, requires that problems have what we call a *well description*, and that this well description satisfies what we call the *finite-subset property*. We define these concepts precisely in Section 2.2, after reviewing preliminaries in Propositional Logic in Section 2.1. Here, we provide an informal description and an illustrative application (which we later formalize): showing that infinite datasets that satisfy SARP are rationalizable.

A *well description* specifies for each problem a (potentially infinite) set of individually finite logical statements over Boolean variables, such that the problem has a solution if and only if there is an assignment of truth values to these Boolean variables under which all these statements hold simultaneously.[3] For example,[4] in a revealed-preferences setting, we can encode a rationalizing preference order using a set of Boolean variables $\{{}_a\mathtt{gt}_b\}$ ("a greater than b"), each being True if $a \succ b$ for the corresponding a, b. We can then express all the required properties of a rationalizing order (completeness, transitivity, antisymmetry, and consistency with whichever outcomes are revealed preferred to others) using logical statements phrased in terms of these variables (infinitely many such statements, but each statement individually finite). An assignment of truth values to the variables under which all statements hold simultaneously corresponds exactly to a solution (a rationalizing order), and vice versa. In particular, for each problem, our set of logical statements has such an assignment if and only if the problem has a solution, and therefore this is a well description.

The preceding well description captures finite and infinite problems equally: the only difference that arises is in the cardinalities of the sets of Boolean variables and logical statements. When the problem is infinitary (e.g., infinite dataset or infinitely many agents), the associated set of logical statements is infinite as well. Yet, each of the logical statements we construct is nonetheless individually finite, that is, it contains only finitely many of the Boolean variables.

Fix a well description, and call the set of logical statements associated with each problem "the description of the problem." The well description satisfies the

[3]The reader may think of an assignment of truth values to the Boolean variables as a "state of the world." In the language of Mathematical Logic, such an assignment under which all statements hold is called a *model for the statements* (perhaps confusingly within an economics setting).

[4]We further elaborate on this example in Section 3.

finite-subset property if every finite subset of the description of any problem belongs to the description of a problem that has a solution. In our example, we identify such a problem that has a solution by restricting the original problem to the data points "mentioned" in the given finite subset. The given finite subset is indeed part of the well description of the restricted problem, and since this problem is finite, it can be solved by known existence results for finite problems, so long as we verify that it "inherits" from the infinite problem any properties required by these results (namely, for our example, satisfying SARP). Lemma 2.5 therefore guarantees the existence of an appropriate solution of the infinite problem.

We prove Lemma 2.5 using *Logical Compactness* (see Section 2.1), a central result in the theory of Propositional Logic. While the above example demonstrates the applicability of our approach to existence results of inherently discrete objects, we also show how to use this approach to scale economic results that go beyond discrete solutions into infinite settings.

2 Framework

In this section, we provide a brief introduction to Propositional Logic (in Section 2.1),[5] and use it to state and prove our main technical lemma (in Section 2.2).

2.1 Propositional Logic Preliminaries

In Propositional Logic, we work with a set of Boolean variables, and study the truth values of statements—called formulae—made up of those variables. We construct formulae by conjoining variables with simple logical operators such as *or*, *not*, and *implies*. Variables are abstract, and do not have meaning on their own—but we can imbue them with "semantic" meaning by introducing formulae that reflect the structure of economic (or other) problems. Once given semantic meaning, the truth or falsity of statements in our Propositional Logic model imply the corresponding results in the associated economic model.

We start by formalizing the idea of *(well-formed propositional) formulae.* To define the set of formulae at our disposal, we first introduce a basic (finite or infinite) set of (Boolean) *variables.* In each section of this paper we introduce a different set of variables built around the economic setting that we model in that section.

Once we have introduced a (finite or infinite) set V of variables, we can define the set of all well-formed formulae inductively:

- 'ϕ' is a well-formed formula for every variable $\phi \in V$.
- '$\neg\phi$' is a well-formed formula for every well-formed formula ϕ.

[5]For a more in-depth look at Propositional Logic primitives and at the Compactness Theorem, see a textbook on Mathematical Logic (e.g., Enderton, 2001; Gonczarowksi and Nisan, 2022). Propositional logic formulae are also used for stating Boolean satisfiability problems.

- ‘$(\phi \vee \psi)$’, ‘$(\phi \wedge \psi)$’, ‘$(\phi \to \psi)$’, and ‘$(\phi \leftrightarrow \psi)$’ are well-formed formulae for every two well-formed formulae ϕ and ψ.

Example 2.1. We could start with a set of four variables $V = \{\mathtt{P}, \mathtt{Q}, \mathtt{R}, \mathtt{S}\}$. Then, each of the following is a well-formed formula:

‘$\mathtt{P}$’ (1) ‘$(\mathtt{P} \vee \mathtt{Q})$’ (2) ‘$\neg(\mathtt{P} \wedge \mathtt{Q})$’ (3) ‘$((\mathtt{P} \wedge \mathtt{R}) \to \mathtt{S})$’ (4)

We sometimes abuse notation by omitting parentheses and writing, e.g., ‘$\phi\vee\psi\vee\xi$’ when any arbitrary placement of parentheses in the formula (e.g., ‘$((\phi \vee \psi) \vee \xi)$’ or ‘$(\phi \vee (\psi \vee \xi))$’) will not make a difference. We sometimes abuse notation even further by writing, e.g., ‘$\bigvee_{i=1}^{10} \phi_i$’ to mean ‘$\phi_1 \vee \phi_2 \vee \cdots \vee \phi_{10}$’ (once again, only when the precise placement of omitted parentheses is of no consequence to our analysis).

We note that while well-formed formulae can be arbitrarily long, each well-formed formula is always finite in length. Thus, for example, a disjunction ‘$\phi_1 \vee \phi_2 \vee \cdots$’ of infinitely many formulae is *not* a well-formed formula. We therefore take special care when we claim that formulae of the form ‘$\bigvee_{\phi\in\Psi} \phi$’ are well-formed, as this is true only if Ψ is finite.

A *model* is a mapping from the set V of all variables to Boolean values, i.e., each variable is mapped either to being True or to being False. This induces a truth value for every formula ‘ϕ’ where $\phi \in V$. A model also induces a *truth value* for all other formulae, defined inductively as follows:

- ‘$\neg\phi$’ is True iff ϕ is False;
- ‘$(\phi \vee \psi)$’ is True iff either or both of ϕ and ψ is True;
- ‘$(\phi \wedge \psi)$’ is True iff both ϕ and ψ are True;
- ‘$(\phi \to \psi)$’ is True iff either ϕ is False or ψ is True or both (that is, ‘$(\phi \to \psi)$’ is False iff both ϕ is True and ψ is False); and
- ‘$(\phi \leftrightarrow \psi)$’ is True iff ϕ and ψ are either both True or both False.

Example 2.2. Given the concept of truth values, we can reinterpret the formulae (1)–(4) as follows:

‘$\mathtt{P}$’	“$\mathtt{P}$ *[*is True*]*”,	(1)
‘$(\mathtt{P} \vee \mathtt{Q})$’	“$\mathtt{P}$ or $\mathtt{Q}$ *[*is True*]*”,	(2)
‘$\neg(\mathtt{P} \wedge \mathtt{Q})$’	“not *(*$\mathtt{P}$ and $\mathtt{Q}$ *[*are both True*])*”,	(3)
‘$((\mathtt{P} \wedge \mathtt{R}) \to \mathtt{S})$’	“$\mathtt{P}$ and $\mathtt{R}$ *[*both being True*]*, implies $\mathtt{S}$ *[*being True*]*”.	(4)

The formula in (2) is True in a model if and only if either ‘$\mathtt{P}$’ or ‘$\mathtt{Q}$’ (or both) are True in that model; the formula in (3) is True in a model unless both ‘$\mathtt{P}$’ and ‘$\mathtt{Q}$’ are True in that model; and the formula in (4) is True in a model unless both ‘$\mathtt{P}$’ and ‘$\mathtt{R}$’ are True in that model while ‘$\mathtt{S}$’ is False in that model.

We say that a formula is *satisfied* by a model if it is True under that model. For example, each of the formulae (1), (2), and (4) is satisfied by the model that assigns value True to all variables, however the formula (3) is not satisfied by that model. We say that a (possibly infinite) set of formulae is *satisfied* by a model if every formula in the set is satisfied by the model. For example, the set of the formulae (1)–(4) is satisfied by the model that assigns value True to all variables except Q. We say that a (possibly infinite) set of formulae is *satisfiable*, or that it *has a model*, if it is satisfied by some model. For example, the set containing 'P' and '¬P' is not satisfiable.

Clearly, if a (finite or infinite) set of formulae Φ is satisfiable, then every subset of Φ is also satisfiable (by the same model), and in particular every *finite* subset of Φ is satisfiable; the *Compactness Theorem for Propositional Logic* gives a surprising and nontrivial converse to this statement.

Theorem 2.3 (Compactness Theorem for Propositional Logic (Gödel, 1930; Malcev, 1936))**.** *A set of formulae Φ is satisfiable if (and only if) every* ***finite*** *subset $\Phi' \subseteq \Phi$ is satisfiable.*

2.2 A Scaling Lemma for Economic Theories

In this section, we use Propositional Logic to derive a sufficient condition for scalability of an economic theorem to infinite cases. This condition, formalized in Lemma 2.5, is at the heart of all of our proofs.

Let $\mathcal{S}$ be a set to which we refer as a set of (potential) *solutions*. Let $\mathcal{P}$ be a set to which we refer as a set of *(economic) problems* whose solutions (if such exist) are in $\mathcal{S}$. For example, for the consumer choice rationalization example from Section 1.1, $\mathcal{S}$ is the set of all consumer preferences over some set of objects, and a problem $P \in \mathcal{P}$ is to rationalize a specific dataset D.

Define $I : \mathcal{P} \times \mathcal{S} \to \{\mathsf{True}, \mathsf{False}\}$ such that $I(P, S)$ is True if and only if S is a solution for P. Note that given a problem P, even if we can easily determine for any given S whether S is a solution of P (i.e., whether $I(P, S)$ is True), it may not be clear just from examining P (and I) whether or not it has *any* solution (i.e., whether there exists $S \in \mathcal{S}$ such that $I(P, S)$ is True). For example, while it is easy to describe when given preferences rationalize a given dataset, it is not immediate from examining a dataset whether there exist preferences that rationalize it. Similarly, while it is easy to describe when a given strategy profile constitutes a Nash equilibrium in a given game, it is not immediate from examining a game whether it admits a Nash equilibrium. Regardless of the economic setting, given a problem, our goal will be to ascertain whether a solution for it indeed exists.

We say that the set $\mathcal{P}$ of problems is a set of *well-describable problems* if for every $P \in \mathcal{P}$ there exists a set Φ_P of well-formed formulae such that P has a solution if and only if Φ_P has a model. We call a collection $(\Phi_P)_{P\in\mathcal{P}}$ of such sets a *well description* of $\mathcal{P}$.

Example 2.4. Consider a set $\mathcal{P}$ of problems where for each problem $P = (X', D) \in \mathcal{P}$, the set of solutions of P consists of all strict preferences over the universe X' that rationalize the given choice data D. To well describe $\mathcal{P}$, one may use the variable of the form ${}_a\mathtt{gt}_b$ from Section 1.1 (where ${}_a\mathtt{gt}_b$ has the semantic interpretation "a is preferred to b"). Specifically, for each $P \in \mathcal{P}$, one may have the set of formulae Φ_P consist of:

1. for all distinct a and b such that in the given choice data, a is chosen from a menu that contains b, the formula '${}_a\mathtt{gt}_b$', requiring that the preferences (that correspond to any model of the formulae) rationalize the given choice data D;
2. for all distinct $a, b \in X'$, the formula '${}_a\mathtt{gt}_b \vee {}_b\mathtt{gt}_a$', requiring that the preferences be complete;
3. for all distinct $a, b \in X'$, the formula '$\neg({}_a\mathtt{gt}_b \wedge {}_b\mathtt{gt}_a)$', requiring that the preferences be antisymmetric;
4. for all distinct $a, b, c \in X'$, the formula '$({}_a\mathtt{gt}_b \wedge {}_b\mathtt{gt}_c) \rightarrow {}_a\mathtt{gt}_c$', requiring that the preferences be transitive.

By construction, $(\Phi_P)_{P\in\mathcal{P}}$ is a well description of $\mathcal{P}$.[6] (The preceding formulae correspond exactly with the formulae (1)–(4) from Section 2.1 upon taking $\mathtt{P} = {}_a\mathtt{gt}_b$, $\mathtt{Q} = {}_b\mathtt{gt}_a$, $\mathtt{R} = {}_b\mathtt{gt}_c$, and $\mathtt{S} = {}_a\mathtt{gt}_c$.)

Given a well description of $\mathcal{P}$, we say that a problem $P \in \mathcal{P}$ satisfies the *finite-subset property* (with respect to the given well description of $\mathcal{P}$) if for every finite subset $\Phi' \subset \Phi_P$ there exists a problem $P' \in \mathcal{P}$ that has a solution and for which $\Phi' \subseteq \Phi_{P'}$. By Theorem 2.3, we then have:

Lemma 2.5 (Scaling Lemma)**.** *Let $\mathcal{P}$ be a set of well-describable problems and let $(\Phi_P)_{P\in\mathcal{P}}$ be a well description of $\mathcal{P}$. Let $P \in \mathcal{P}$. If P satisfies the finite-subset property, then P has a solution.*

Proof. Let $P \in \mathcal{P}$ be a problem satisfying the finite-subset property. By well describability, it is enough to show that Φ_P is satisfiable. By Theorem 2.3, it is therefore enough to show that every finite $\Phi' \subset \Phi_P$ is satisfiable. Let Φ' be such a finite subset. Since P satisfies the finite-subset property, there exists $P' \in \mathcal{P}$ that has a solution such that $\Phi' \subseteq \Phi_{P'}$. Since P' has a solution, by our well-describability assumption, we have that $\Phi_{P'}$ is satisfiable by some model. Since $\Phi' \subseteq \Phi_{P'}$, the same model also satisfies Φ', and so Φ' is satisfiable as required. □

[6]In fact, an even stronger property holds: the set of solutions of P is in *one-to-one* correspondence with the set of models of Φ_P (see Section 3 for more details on this correspondence). While such a one-to-one correspondence holds in many of our applications, this is not required for our arguments.

As demonstrated in Section 1.1, in many cases of interest, the existence of a solution for an appropriate P' for any Φ' can be established by finite-case theorems (for example, on rationalizability of finite datasets or stable matching in finite markets). Thus by Lemma 2.5 we obtain existence of solutions for the infinite case of such problems as well. In this paper, we demonstrate the wide applicability of Lemma 2.5 to a wide range of economic problems.

3 Warm Up: Rational Choice Functions

We begin with a classic revealed preference setup. Let X be a (possibly infinite) set of goods. The set of *menus* includes all finite subsets of X. A *dataset* $D \subseteq \{(S,a) \in 2^X \times X \mid a \in S\}$ consists of the (unique) respective choices made by an agent in a (possibly infinite) subset of menus. A pair $(S,a) \in D$ is interpreted to mean that the agent selected $a \in S$ when presented with the menu S. We say that a dataset D is *rationalized* by a strict preference relation $\succ$ (complete, antisymmetric, and transitive) over X, if for every $(S,a) \in D$, the agent's choice a is the maximal element from S according to $\succ$. A dataset is *rationalizable* if it is rationalized by some strict preference relation over X.

Given a dataset and a pair of goods, $x, y \in X$, we say that x is *revealed preferred* to y ($x \succ^R y$) if x is chosen from a menu that includes y. A dataset satisfies the *strong axiom of revealed preferences (SARP)* if $\succ^R$ is acyclic (i.e., there does not exist $k > 1$ and $x_1, \ldots, x_k \in X$ such that $x_i \succ^R x_{i+1}$ for all i, and in addition $x_k \succ^R x_1$). Clearly, satisfying SARP is a necessary condition for a dataset to be rationalizable. A classic result due to Richter (1966) and Hansson (1968) is that satisfying SARP is also sufficient. Theorem 3.1 states this result for the special case of a finite dataset.

Theorem 3.1. *A finite dataset is rationalizable if and only if it satisfies SARP.*

The general version of Theorem 3.1, due to Richter (1966) and Hansson (1968), builds on Szpilrajn' Extension Theorem, a fundamental result whose variants are used to prove many key results in the theory of revealed preferences (Chambers and Echenique, 2009). Ok (2007, p. 17) explains that "[a]lthough it is possible to prove this [fundamental result of order theory] by mathematical induction when X is finite,[7] the proof in the general case is built on a relatively advanced method[...]." Indeed, the standard way to prove an infinite-dataset version of Theorem 3.1 (equivalently, to prove Szpilrajn's Extension Theorem) as well as to prove variant results, is to use Zorn's Lemma (see, e.g., Richter, 1966; Duggan, 1999; Mas-Colell, Whinston, and Green, 1995, Proposition 3.J.1; Chambers and

[7] I.e., the setting corresponding to Theorem 3.1; for an explicit proof see, e.g., Lahiri (2002).

Echenique, 2016, Theorems 1.4 and 1.5).[8]

As a warm-up, we use Lemma 2.5 to scale Theorem 3.1 to also apply to infinite datasets (implicitly reproving Szpilrajn's Extension Theorem). While Lemma 2.5 relies on Logical Compactness, which, like Zorn's Lemma, relies on some variant of the Axiom of Choice,[9] we suspect that the proof we present here may complement the standard approach. In particular, our argument may in some ways be more accessible to students than the traditional proof because it avoids the "overhead" of understanding the full statement of Zorn's Lemma.[10]

Theorem 3.2 (Richter, 1966; Hansson, 1968)**.** *A (possibly infinite) dataset is rationalizable if and only if it satisfies SARP.*

Proof of Theorem 3.2. As noted, the "only if" direction is immediate, so we prove the "if" direction using Lemma 2.5.

Definition of $\mathcal{P}$: Fixing X, let $\mathcal{P}$ be the set of all pairs (X', D) such that $X' \subseteq X$ and D is a dataset with menus over goods in X' that satisfies SARP. A *solution* for a pair $(X', D) \in \mathcal{P}$ is a strict preference order over X' that rationalizes D.

Well describability: We introduce a variable ${}_a\mathtt{gt}_b$ for every pair of distinct $a, b \in X$. In what follows, for each $(X', D) \in \mathcal{P}$ we define a set $\Phi_{(X',D)}$ of formulae over these variables so that models (over the variables that appear in $\Phi_{(X',D)}$) of $\Phi_{(X',D)}$ are in one-to-one correspondence with the (not-yet-proven-to-be-nonempty) set of solutions for (X', D). The correspondence is obtained by endowing the variable ${}_a\mathtt{gt}_b$ with the semantic interpretation "a is preferred to b." That is, it maps a model for $\Phi_{(X',D)}$ to the preference $\succ$ such that for every distinct $a, b \in X'$, we have that $a \succ b$ if and only if the variable ${}_a\mathtt{gt}_b$ is $\mathtt{True}$ in that model. We define the set $\Phi_{(X',D)}$ to consist of the formulae from Example 2.4.

Finite-subset property: Let $(\bar{X}, D) \in \mathcal{P}$. Let $\Phi' \subset \Phi_{(\bar{X},D)}$ be a finite subset. Since Φ' is finite, it "mentions" (through variables used) only finitely many elements of $\bar{X}$; denote the set of these elements by $X' \subset \bar{X}$. Let

$$D' \triangleq \big\{(S \cap X', a) \;\big|\; (S, a) \in D \text{ and } a \in X'\big\}.$$

By definition, $\Phi' \subseteq \Phi_{(X',D')}$. Furthermore, D' satisfies SARP since had a cycle been induced by D', it would have also been induced by D, however D satisfies SARP, so it induces no cycles. Hence, by Theorem 3.1, D' is rationalizable by some strict preference order over X'. Therefore, $(\bar{X}, D)$ satisfies the finite-subset property. Thus, by Lemma 2.5, D is rationalizable by a strict preference order over $\bar{X}$. □

[8]Mandler (2020) provides an alternative simple proof of Richter and Hansson's result.

[9]Strictly speaking, under ZF set theory, Logical Compactness is weaker than Choice.

[10]Mas-Colell, Whinston, and Green (1995), similarly to Ok (2007), label their proof (which uses Zorn's Lemma) as "advanced."

4 Applicability of the Same Proof to Other Settings: Limited Attention

A notable strength of our approach is that it is agnostic to the methods used to prove the finite result being scaled. Therefore, the same argument can be used to scale similar statements even when the finite-case proofs of these statements hinge on very different tools. As an illustration, we use essentially the same proof as in Section 3 to scale the WARP-limited attention (WARP-LA) result of Masatlioglu et al. (2012) to infinite datasets, despite the Masatlioglu et al. proof using approaches that are quite different from those used to prove the result of Theorem 3.1.

The setup resembles that of Section 3: let X be a (possibly infinite) set of goods. A *menu* is a finite nonempty subset of X. A dataset is *full* if it consists of the (unique) respective choices made by an agent in *all possible* (infinitely many, if X is infinite) menus. Unlike in Section 3, in this setting a rational agent might fail to choose her favorite object from a menu if this object does not capture her attention. The objects that capture the agent's attention in each menu are described by an *attention filter*. Formally, a *filter* is a function Γ that maps each menu S to a menu $\Gamma(S) \subseteq S$. A filter is an *attention filter* if $\Gamma\big(S \setminus \{x\}\big) = \Gamma(S)$ for every menu S and $x \in S \setminus \Gamma(S)$. A dataset is *limited-attention rationalizable* if there exist a strict preference relation $\succ$ and an attention filter Γ such that for every menu S in the dataset, the agent's choice is the most-preferred element in $\Gamma(S)$ according to $\succ$.

Analogously to Theorem 3.1, Masatlioglu et al. (2012) uncover a condition, WARP-LA,[11] that is necessary and sufficient for limited-attention rationalizability:

Theorem 4.1 (Masatlioglu et al., 2012)**.** *A full finite dataset is limited-attention rationalizable if and only if it satisfies WARP-LA.*

As is the case for SARP, if a certain dataset satisfies WARP-LA, then so does any sub-dataset. This suffices for us to scale Theorem 4.1 to infinite datasets using the same approach we used in Section 3.

Theorem 4.2. *A full (possibly infinite) dataset is limited-attention rationalizable if and only if it satisfies WARP-LA.*

The well description that we build to prove Theorem 4.2 is conceptually similar to the one from our proof of Theorem 3.2, but requires slight modification due to the addition of the attention filter. The idea is to have, as before, for every pair $a, b \in X$ a variable ${}_a\mathtt{gt}_b$ that will be True in a model if and only if a precedes b in the preference relation corresponding to the model. But furthermore, for every

[11] A full dataset satisfies WARP-LA if, for any menu S, there exists $x^* \in S$ such that for any menu T that includes x^*, if $(T, x) \in D$ for some $x \in S$ and $(T \setminus \{x^*\}, x) \notin D$, then $(T, x^*) \in D$.

menu S and $\emptyset \neq T \subseteq S$, we introduce a variable $\texttt{attn}_{(S,T)}$ that will be True in a model if and only if $\Gamma(S) = T$ for the attention filter corresponding to the model, and upon whose value the formulae that represent the dataset observations will be conditioned. So, if a is chosen from S in the dataset, for every $b \in S \setminus \{a\}$ instead of having a single formula ${}_a\texttt{gt}_b$ mandating that a precede b in the preference relation, we have for each $b \in T \subseteq S$ a formula that says "if $\Gamma(S) = T$ then a precedes b in the preference relation," i.e., '$\texttt{attn}_{(S,T)} \rightarrow {}_a\texttt{gt}_b$'. Additionally, we introduce formulae that say that for each menu S, there exists precisely one T such that $\texttt{attn}_{(S,T)}$ holds. Finally, for each menu S, each $T \subset S$ and $x \in S \setminus T$ we introduce the formula '$\texttt{attn}_{(S,T)} \rightarrow \texttt{attn}_{(S\setminus\{x\},T)}$', requiring that the filter be an attention filter. Except for these additions, the proof runs along the same lines as that of Theorem 3.2; we relegate the details to Appendix A.

4.1 Conditional Scaling

Limited-attention rationalizability, as defined above, flexibly accommodates a wide array of attention filters. But, in some cases, one may wish to impose additional structure (e.g., requiring that the agent always pays attention to at least two options), or to consider filters that fall outside the domain of attention filters (Masatlioglu et al., 2012, discuss a wide array of examples from the literature). In many cases of interest, the constraint on the filters to be considered takes the form $\forall S_1 \forall S_2 \cdots \forall S_n : H\big(S_1, \Gamma(S_1), S_2, \Gamma(S_2), \ldots, S_n, \Gamma(S_n)\big) = \mathsf{True}$ for some predicate H that takes $2n$ menus, where all S_i are menus.[12] Such restrictions can be encoded into our well description (in the same fashion that we encoded the restriction that Γ is an attention filter in our proof of Theorem 4.2).[13]

This provides an opportunity to point out that our framework is agnostic not only to *how* the finite-case theorem being scaled was proven (as already discussed), but furthermore, to *whether* it has even been proven. Indeed, even absent a proof for the finite-case theorem, our framework can yield conditional statements. Consider a result (potentially, one that could be uncovered in the future) that, for some predicate H, determines that a finite dataset is H-rationalizable (i.e., rationalizable using a filter that meets the above requirement with respect to H) if and only if it satisfied some given condition WARP-H defined on finite datasets. Extend the definition of WARP-H to infinite datasets by defining that an infinite dataset satisfies WARP-H if and only if every finite sub-dataset of it satisfies

[12]For example, requiring that the agent always pays attention to at least two options can be expressed with $n = 1$, setting $H(S_1, T_1)$ to be True unless $|S_1| > 1$ yet $|T_1| = 1$. As another example, the filter being an attention filter could have been expressed in this form with $n = 2$, setting $H(S_1, T_1, S_2, T_2)$ to be True unless $S_1 = S_2 \cup \{x\}$, $x \notin T_1$, and yet $T_1 \neq T_2$.

[13]The only change to the well description would be to add one more (finite) formula type, for every $S_1, \ldots, S_n$: '$\bigvee_{T_1 \subseteq S_1, \ldots, T_n \subseteq S_n : H(S_1, T_1, \ldots, S_n, T_n)} \bigwedge_{i=1}^{n} \texttt{attn}_{(S_i, T_i)}$'.

WARP-H.[14] A proof completely analogous to our proof of Theorem 4.2 then also proves it imediately applies to infinite datasets:

Theorem 4.3. *If it holds that a finite dataset is H-ratiolalizable iff it satisfies WARP-H, then it also holds that an infinite dataset is H-ratiolalizable iff it satisfies WARP-H.*

5 Handling Nondiscrete Solution Concepts: Rationalizing Consumer Demand

We now move to rationalizing consumption behavior in the presence of prices. For the most part, in this section we follow the notation of Reny (2015). Fix a number of goods $m \in \mathbb{N}$. A dataset $D \subset \left(\mathbb{R}^m_+ \setminus \{\bar{0}\}\right) \times \mathbb{R}^m_+$ with generic element $(\bar{p}, \bar{x}) \in D$ represents a set of observations, where in each, a consumer with a budget faces a price vector $\bar{p} \neq \bar{0}$ and chooses to consume the bundle $\bar{x}$. A utility function $u : \mathbb{R}^m_+ \to \mathbb{R}$ *rationalizes* the dataset D if for every $(\bar{p}, \bar{x}) \in D$ and every $\bar{y} \in \mathbb{R}^m_+$, it holds that if $\bar{p} \cdot \bar{y} \leq \bar{p} \cdot \bar{x}$ (i.e., $\bar{y}$ can also be bought with the budget) then $u(\bar{y}) \leq u(\bar{x})$, and if $\bar{p} \cdot \bar{y} < \bar{p} \cdot \bar{x}$ (i.e., $\bar{y}$ can be bought without spending the entire budget) then $u(\bar{y}) < u(\bar{x})$.[15] If only the former implication holds for every such $(\bar{p}, \bar{x})$ and $\bar{y}$, then we say that u *weakly rationalizes* D.

A dataset D satisfies the *Generalized Axiom of Revealed Preference (GARP)* if for every (finite) sequence $(\bar{p}_1, \bar{x}_1), \ldots, (\bar{p}_k, \bar{x}_k) \in D$, if for every $i \in \{1, 2, \ldots, k{-}1\}$ it holds that $\bar{p}_i \cdot \bar{x}_{i+1} \leq \bar{p}_i \cdot \bar{x}_i$, then $\bar{p}_k \cdot \bar{x}_1 \geq \bar{p}_k \cdot \bar{x}_k$. It is straightforward from the definitions that satisfying GARP is a precondition for rationalizability (indeed, otherwise we would have that $u(\bar{x}_1) \geq u(\bar{x}_2) \geq \cdots \geq u(\bar{x}_k) > u(\bar{x}_1)$ for any rationalizing utility function u). In a celebrated result, Afriat (1967) showed that GARP is also a sufficient condition for rationalizability of a finite dataset—and furthermore GARP is a sufficient condition for rationalizability of such a dataset by a utility function with many properties that are often assumed in simple economic models. This finding implies that the standard economic model of consumer choice has no testable implications beyond GARP.

Theorem 5.1 (Afriat, 1967). *A finite dataset $D \subseteq \left(\mathbb{R}^m_+ \setminus \{\bar{0}\}\right) \times \mathbb{R}^m_+$ satisfies GARP if and only if it is rationalizable. Moreover, when GARP holds, there exists a utility function that rationalizes D that is continuous, concave, nondecreasing, and strictly increasing when all coordinates strictly increase.*

[14] If for finite datasets, WARP-H is a "no cycles of some form" condition (like SARP, WARP, or WARP-LA), then extending it to infinite datasets results in the same condition for infinite datasets as well: no (finite) cycles of that form.

[15] This assumption rules out trivial rationalizations such as constant utility functions. See Chambers and Echenique (2016) for a more detailed discussion.

There are well-known examples of infinite datasets that are generated by quasiconcave utility functions but may not be rationalized by a concave utility function (see Aumann, 1975; Reny, 2013). Kannai (2004) and Apartsin and Kannai (2006) provide necessary conditions, stronger than GARP, for rationalizability by a concave function. Recently, Reny (2015) unified the literature and clarified the boundaries of Afriat's theorem by showing that GARP is indeed necessary and sufficient for rationalization of even infinite datasets—and in fact, GARP also guarantees rationalizability by a utility function with many desired properties (yet not all the properties that are attainable in the finite case).

Theorem 5.2 (Reny, 2015)**.** *A (possibly infinite) dataset $D \subseteq \left(\mathbb{R}^m_+ \setminus \{\bar{0}\}\right) \times \mathbb{R}^m_+$ satisfies GARP if and only if it is rationalizable. Moreover, when GARP holds, there exists a utility function that rationalizes D that is quasiconcave, nondecreasing, and strictly increasing when all coordinates strictly increase.*

Reny (2015) provided examples showing that continuity and concavity (the properties of the rationalizing utility function from Theorem 5.1 that are absent from Theorem 5.2) cannot be guaranteed to be attainable for any rationalizable dataset. Reny (2015) then proved Theorem 5.2 using a novel construction that—unlike Afriat's construction—applies also to infinite data sets. We instead give a concise alternative proof of Theorem 5.2 by scaling Theorem 5.1 as a black box using Lemma 2.5.[16] One of the challenges in our argument is that utility functions have an (uncountably) infinite range, so it is not *a priori* obvious how to encode such a function by a model defined via individually finite formulae (e.g., how to require that each bundle is associated with some real number that represents the utility from it); to overcome this challenge, our well description encodes the utility from each bundle as the limit of a sequence of discrete utilities. This approach, in turn, introduces additional challenges, such as how to make sure, using only constraints on these discrete functions, that the limit utilities satisfy all desired properties. This is particularly challenging with properties that are not preserved by limits, such as being strictly increasing.

Proof of Theorem 5.2. As noted, the "only if" direction is immediate, so we prove the "if" direction using Lemma 2.5.

<u>Definition of $\mathcal{P}$:</u> Fixing m, let $\mathcal{P}$ be the set of all datasets $D \subseteq \left(\mathbb{R}^m_+ \setminus \{\bar{0}\}\right) \times \mathbb{R}^m_+$ satisfying GARP. A *solution* for a dataset $D \in \mathcal{P}$ is a utility function that ra-

[16] While we prove the result of Reny (2015) in its full generality, it is worth noting that Reny's proof does not use the Axiom of Choice, while ours does to some extent. More specifically, the Compactness Theorem, which we use for proving Lemma 2.5, is equivalent (under ZF) to the Boolean Prime Ideal (BPI) Theorem (equivalently, to the Ultrafilter Lemma), which is known to be a "weaker form of the Axiom of Choice" in the sense that ZF+BPI is strictly weaker than ZFC but strictly stronger than ZF (see, e.g., Halbeisen, 2017, Theorems 6.7 and 8.16).

tionalizes D that is quasiconcave, nondecreasing, and strictly increasing when all coordinates strictly increase.

Well describability: We set $\varepsilon_n \triangleq 2^{-n}$ for every $n \in \mathbb{N}$. We introduce a variable $\texttt{utility}^n_{\bar{x},v}$ for every $n \in \mathbb{N}$, every $\bar{x} \in \mathbb{R}^m_+$, and every $v \in V_n \triangleq \{0, \varepsilon_n, 2\cdot\varepsilon_n, \ldots, 1\}$. In what follows, for each $D \in \mathcal{P}$ we define a set Φ_D of formulae over these variables so that models of Φ_D are in one-to-one correspondence with the (not-yet-proven-to-be-nonempty) set of solutions for D that satisfy certain properties (we then have to show that the existence of any solution implies the existence of a solution with such properties). The correspondence is obtained by endowing the variable $\texttt{utility}^n_{\bar{x},v}$ with the semantic interpretation "$\lfloor u(\bar{x}) \rfloor_{\varepsilon_n} = v$ for the corresponding utility function u," where for every $n \in \mathbb{N}$ and every x we denote by $\lfloor x \rfloor_{\varepsilon_n} \triangleq 2^{-n} \cdot \lfloor 2^n \cdot x \rfloor$ the rounding-down of x to the nearest multiple of ε_n. Fixing an enumeration $(\bar{q}^k_1, \bar{q}^k_2)_{k=1}^{\infty}$ of the countable set $\{(\bar{q}_1, \bar{q}_2) \in \mathbb{Q}^m \times \mathbb{Q}^m \mid \bar{q}_1 \ll \bar{q}_2\}$,[17] we define the set Φ_D to consist of the following formulae:

1. for all $n \in \mathbb{N}$ and all $\bar{x} \in \mathbb{R}^m_+$, the (finite!) formula '$\bigvee_{v \in V_n} \texttt{utility}^n_{\bar{x},v}$', requiring that $\bar{x}$ have a rounded-down-to-ε_n utility in $[0,1]$;
2. for all $n \in \mathbb{N}$, all $\bar{x} \in \mathbb{R}^m_+$, and all distinct $v, w \in V_n$, the formula '$\texttt{utility}^n_{\bar{x},v} \to \neg\texttt{utility}^n_{\bar{x},w}$', requiring that the rounded-down-to-ε_n utility from $\bar{x}$ be unique;
3. for all $n \in \mathbb{N}$, all $\bar{x} \in \mathbb{R}^m_+$, and all $v \in V_n$, the formula '$\texttt{utility}^n_{\bar{x},v} \to \big(\texttt{utility}^{n+1}_{\bar{x},v} \vee \texttt{utility}^{n+1}_{\bar{x},v+\varepsilon_{n+1}}\big)$', requiring that $\lfloor u(\bar{x}) \rfloor_{\varepsilon_n} = \lfloor \lfloor u(\bar{x}) \rfloor_{\varepsilon_{n+1}} \rfloor_{\varepsilon_n}$;
4. for all $n \in \mathbb{N}$, all $\bar{x}, \bar{y} \in \mathbb{R}^m_+$, all convex combinations $\bar{z} \in \mathbb{R}^m_+$ of $\bar{x}, \bar{y}$, and all $v, w \in V_n$, the (finite) formula '$\big(\texttt{utility}^n_{\bar{x},v} \wedge \texttt{utility}^n_{\bar{y},w}\big) \to \bigvee_{v' \in V_n : v' \geq \min\{v,w\}} \texttt{utility}^n_{\bar{z},v'}$', requiring that the rounded-down-to-ε_n utility function be quasiconcave;
5. for all $n \in \mathbb{N}$, all $\bar{x}, \bar{y} \in \mathbb{R}^m_+$ s.t. $\bar{x} \leq \bar{y}$, and all $v \in V_n$, the (finite) formula '$\texttt{utility}^n_{\bar{x},v} \to \bigvee_{w \in V_n : w \geq v} \texttt{utility}^n_{\bar{y},w}$', requiring that the rounded-down-to-ε_n utility function be nondecreasing;
6. for all $k \in \mathbb{N}$ and all $n > k$, the (finite) formula '$\texttt{utility}^n_{\bar{q}^k_1,v} \to \bigvee_{w \in V_n : w \geq v + 2^{-k-1}} \texttt{utility}^n_{\bar{q}^k_2,w}$', requiring that starting at some n, the rounded-down-to-ε_n utility from $\bar{q}^k_2$ be greater by at least 2^{-k-1} than the rounded-down-to-ε_n utility from $\bar{q}^k_1$;
7. for all $n \in \mathbb{N}$, all datapoints $(\bar{p}, \bar{x}) \in D$, all $\bar{y} \in \mathbb{R}^m_+$ s.t. $\bar{p} \cdot \bar{y} \leq \bar{p} \cdot \bar{x}$, and all $v \in V_n$, the (finite) formula '$\texttt{utility}^n_{\bar{x},v} \to \bigvee_{w \in V_n : w \leq v} \texttt{utility}^n_{\bar{y},w}$', requiring that the rounded-down-to-ε_n utility weakly rationalize D.

We now argue that $(\Phi_D)_{D \in \mathcal{P}}$ is a well description of $\mathcal{P}$. Let $D \in \mathcal{P}$.

We first claim that every model that satisfies Φ_D corresponds to a solution for D. Fix a model for Φ_D. For every $\bar{x} \in \mathbb{R}^m_+$ and every $n \in \mathbb{N}$, let $v_n \in V_n$ be the value

[17] For $\bar{x}, \bar{y} \in \mathbb{R}^m$, we write $\bar{x} \ll \bar{y}$ to denote that $x_i < y_i$ for every $i = 1, \ldots, m$.

such that $\texttt{utility}^n_{\bar{x},v_n}$ is True in the model (well defined by the first and second formula-types above), and define $u(\bar{x}) = \lim_{n\to\infty} v_n$ (well defined, e.g., by the third formula-type above since v_n is a Cauchy sequence). The resulting utility function u is a limit of nondecreasing quasiconcave functions (by the fourth and fifth formula-types above) that weakly rationalize the data (by the seventh formula-type above). Hence, u itself is a nondecreasing quasiconcave function that weakly rationalizes the data. Furthermore, for every $\bar{x}, \bar{y} \in \mathbb{R}^m_+$ s.t. $\bar{x} \ll \bar{y}$, there exist two rational number vectors "in between" them, i.e., there exists $k \in \mathbb{N}$ s.t. $\bar{x} \ll \bar{q}^k_1 \ll \bar{q}^k_2 \ll \bar{y}$. Therefore, we have that $u(\bar{x}) \leq u(\bar{q}^k_1) \leq u(\bar{q}^k_2) - 2^{-k-1} < u(\bar{q}^k_2) \leq u(\bar{y})$ (the second inequality stems from this inequality holding for almost all functions of which u is the limit, by the sixth formula-type above), so u is strictly increasing when all coordinates strictly increase. Finally, since u weakly rationalizes D and is also strictly increasing when all coordinates strictly increase, then u also rationalizes D.

Second, we claim that if D has a solution, then Φ_D has a model. Fix a solution u for D, and let $\bar{u}(\bar{x}) \triangleq 1/4 + (1/2\pi) \cdot \arctan(u(\bar{x})) + \sum_{k:\bar{q}^k_2 \leq \bar{x}} 2^{-k-1}$ for every $\bar{x} \in \mathbb{R}^m_+$. As this transformation of utilities is strictly monotone, the resulting function $\bar{u}$ still rationalizes the data, and is quasiconcave, nondecreasing, and strictly increasing when all coordinates strictly increase. Furthermore, the sum of the first two summands is in $[0, 1/2]$, and so is the third summand, so the overall sum is in $[0, 1]$. Finally, due to the third summand, $\bar{u}(\bar{q}^k_2) > \bar{u}(\bar{q}^k_1) + 2^{-k-1}$ for every $k \in \mathbb{N}$. Using $\bar{u}$ we can therefore construct a model for Φ_D (by setting each $\texttt{utility}^n_{\bar{x},v}$ to be True iff $v = \lfloor \bar{u}(x) \rfloor_{\varepsilon_n}$), and so Φ_D has a model. To sum up, $(\Phi_D)_{D\in\mathcal{P}}$ is a well description of $\mathcal{P}$.

Finite-subset property: Let $D \in \mathcal{P}$. Let $\Phi' \subset \Phi_D$ be a finite subset. Since Φ' is finite, there are only finitely many formulae of the above seventh type (the only formula type that depends on the dataset) in Φ'. Let $D' \subset D$ be the set of datapoints that induce these formulae. By definition, $\Phi' \subseteq \Phi_{D'}$. Furthermore, D' satisfies GARP since any sub-dataset of D satisfies GARP, and hence, by Theorem 5.1, D' is rationalizable. Therefore, D satisfies the finite-subset property. Thus, by Lemma 2.5, D is rationalizable. □

A natural question is why the same argument cannot be used to scale Theorem 5.1 while maintaining concavity rather than quasiconcavity. The short answer is that—due to the requirement that u be strictly monotone, and the inherent need to make each formula finite—our proof of Theorem 5.2 relies heavily on the fact that quasiconcavity, unlike concavity, is maintained under weakly monotone transformations (such as the mapping of u to $\bar{u}$); we discuss this further in Section 6.

5.1 Additional Application of the Same Proof: Rational Inattention

Our proof of Theorem 5.2 is quite a bit more flexible than one might imagine. In Appendix B, we use essentially the same well description to scale—from finite to infinite datasets—the seminal result of Caplin and Dean (2015) in quite a different rationalization domain: a state-dependent stochastic choice dataset has a costly information acquisition representation if and only if it satisfies the *No Improving Action Switches (NIAS)* and *No Improving Attention Cycles (NIAC)* conditions. Caplin et al. (2017) recently proved the infinite version of this result via a novel proof that diverges from Caplin and Dean's proof of the finite case.[18] We reprove this result using essentially the same well description as in our proof of Theorem 5.2, despite the differences between the two settings considered, and despite the fact that neither the original proofs of the finite versions nor the original proofs of the infinite versions of any of these quite different theorems share any common core technique. The main difference between the two well descriptions is that this application does not require strict monotonicity. Therefore, the sixth formula-type of the above well description is not required, and the proof that a solution implies a model is simpler as it does not require carefully "massaging" the function u into $\bar{u}$ as above.

6 Remarks and Limitations

In the preceding sections, we demonstrated the versatility of our approach across several revealed-preference settings. Our approach can be used to scale many additional finite-data results to encompass infinite datasets. For example, it may allow us to scale results such as those of Cattaneo et al. (2020) on the existence of random attention representation, or Filiz-Ozbay and Masatlioglu (2020) on progressive random choice.[19] In addition to scaling a wide array of finite-data results to encompass infinite datasets, our approach can also be used to adapt finite-data rationalization results to support parametric restrictions, as in Hu et al. (2021), since such restrictions often translate into infinitely many constraints. Our approach, however, is not without limitations. In this section, we provide some remarks on limitations of our proofs within revealed preference. A more high-level discussion of settings in which our approach is not applicable is provided in Section 7.2.

Afriat's theorem (Theorem 5.1) guarantees that finite demand datasets satisfying GARP can be rationalized using a concave utility function. But, there are well-known examples of quasiconcave utility functions whose full (infinite)

[18] de Oliveira et al. (2017) provide a similar result for infinite datasets of a different kind.

[19] We thank Yusufcan Masatlioglu for proposing these applications.

demand dataset (which satisfies GARP since it is derived from the choices of a utility function) cannot be rationalized using a concave utility function. In Section 5, we reproved the main result of Reny (2015) that unified these settings: any demand dataset, finite or infinite, that satisfies GARP can be rationalized using a quasiconcave utility function (Theorem 5.2).

By Lemma 2.5, the existence of a counterexample, together with the correctness of Afriat's theorem for finite datasets, implies that the existence of a concave rationalizing utility function (as guaranteed by Afriat's theorem) has no well description that satisfies the finite-subset property. This might seem puzzling since a simple modification to the fourth formula type in our proof (which imposes quasiconcavity) can be used to impose concavity (as in our similar scaling proof in Appendix B), and so should seemingly result in a well description as required. The answer to this puzzle is that this well description does not, in fact, satisfy the finite-subset property. Specifically, the sixth formula type in our proof makes a stronger monotonicity requirement than the monotonicity that is guaranteed by Afriat's theorem, and therefore Afriat's theorem cannot be used to show that the finite-subset property required by Lemma 2.5 holds.

To address this issue, a natural approach would be to change the monotonicity requirement that we use to require only strict monotonicity, as guaranteed by Afriat's theorem. But, it is not possible to well-describe strict monotonicity with our variables (since strict inequalities are not preserved in the limit). Our way around this limitation was to make a stronger requirement that is well-describable. But in order to use Afriat's theorem to show that the finite-subset property holds, we had to relax the concavity requirement (recall that our proof applied monotonic transformations to the utility function; while these transformations do not preserve concavity, they do preserve quasiconcavity). We note that while this may appear to be an artefact of using Lemma 2.5, the existence of the abovementioned counterexample guarantees that no other approach could circumvent this issue. The tradeoff between strengthening monotonicity and weakening concavity, so that well describability and the finite-subset property are satisfied, sheds some new light on what breaks in the infinite case, which at first glance might look like an issue with concavity, but at a deeper look reveals itself as an issue with strict monotonicity. This affords some degree of intuition for "why" the concavity assumption in Theorem 5.1 must be relaxed to quasiconcavity when scaling it to infinite datasets.

We note that when we use a very similar proof in Appendix B to scale a result by Caplin and Dean (2015), we do require concavity rather than merely quasiconcavity. This is possible because in that result only weak (rather than strict) monotonicity (in information) is required, which can be well described without being strengthened. Contrasting these two proofs provides yet another example of the power of our approach to very tangibly pinpoint why certain conditions can be maintained when some theorems are scaled but not when others are.

Meanwhile, Theorem 5.2 illustrates some of the limitations of our framework. A Propositional Logic formulation precludes the use of quantifiers (e.g., "*there exists* a positive gap by which the utility from $\bar{q}_2$ is greater than the utility from $\bar{q}_1$"), and also precludes infinitely long formulae such as infinite disjunctions (e.g., "the utility from $\bar{q}_2$ is greater than the utility from $\bar{q}_1$ by at least *one of the following infinitely many* positive gaps"). This prohibits the well description of certain properties of interest (e.g., strict monotonicity, unless strengthened) without the use of variables that refer to infinitely many objects. But such variables oftentimes hinder the ability to invoke finite theorems to show that the finite-subset property holds.

The requirement of strict monotonicity underlies another well-known counterexample. While a strict preference order over a finite set of objects can always be represented by a utility function, the same need not be true when the set of objects is uncountable.[20] Accordingly, any attempt to use our strengthened monotonicity requirement to scale the finite case is of course bound to fail when the set of objects is uncountable. It is instructive to consider how it would fail. Recall that our strengthened monotonicity requires fixed positive gaps between various utility values. In the case of Afriat's theorem, requiring countably many such gaps sufficed, and hence the required gap lengths could be chosen so that their sum is finite. By contrast, scaling the existence of a utility representation for strict preferences to uncountable sets would involve requiring uncountably many positive gaps. This means that the sum of lengths of required gaps would be infinite, and so some objects would not be associated with a finite utility.

7 Beyond Revealed Preferences

Our approach is not in any way limited to revealed preferences. In this section, we illustrate its applicability in another domain: non-cooperative game theory. We furthermore provide examples where our framework is inapplicable.[21]

7.1 Nash Equilibria in Games on Infinite Graphs

In this section, we turn to the setting of games on graphs (see, e.g., Kearns, 2007, and the references therein),which includes overlapping-generations models, even with infinite time. We use Lemma 2.5 to show the existence of a Nash equilibrium in games on infinite graphs. Our result here is covered by Peleg (1969) (who directly scales the seminal existence result of Nash, 1951), but we give a new proof

[20]For example, lexicographic preferences over $\mathbb{R}^2$ or any strict preference order over $2^{\mathbb{R}}$.

[21]Our working paper (Gonczarowski, Kominers, and Shorrer, 2023) includes applications in other domains.

that uses the same principled approach we use throughout this paper.

Here, we use Lemma 2.5 to scale the existence of arbitrarily good approximate Nash equilibria, and then show that the existence of such approximate equilibria implies the existence of an exact equilibrium. This two-step proof strategy is chosen for convenience: with additional variables, it is easy to encode the second step of the proof into the logical formulation just like we did in Section 5.[22]

In a *game on a graph*, there is a (potentially infinite) set of players I, each having a finite set of pure *strategies* S_i. Each player $i \in I$ is linked to a finite set of neighbors $N(i) \subset I$ with $i \in N(i)$, and her utility only depends on the strategies played by players in the set $N(i)$.[23] This setting occurs, for example, in infinite-horizon overlapping-generations models, where at each point in time there are only finitely many players alive, and a player's utility depends only on the behavior of contemporary players. For any player i we denote by $\Sigma_i \triangleq \Delta(S_i)$ the set of *mixed strategies* (i.e., distributions over pure strategies) of player i. A *mixed-strategy profile* $(\sigma_i)_{i\in I}$ is a specification of a mixed strategy $\sigma_i \in \Sigma_i$ for every player $i \in I$. A mixed-strategy profile $(\sigma_i)_{i\in I}$ is a *Nash equilibrium* if for every $i \in I$ and every possible deviating strategy $\sigma'_i \in \Sigma_i$, it holds that $u_i(\sigma_{N(i)}) \geq u_i(\sigma'_i, \sigma_{N(i)\setminus\{i\}})$.

Games on *finite* graphs have finitely many players and finitely many strategies per player; hence, the seminal analysis of Nash (1951) implies that they have Nash equilibria.

Theorem 7.1 (Follows from Nash (1951)). *Every game on a finite graph has a Nash equilibrium.*

Our main result of this section is that Nash equilibria are guaranteed to exist even in games on infinite graphs.

Theorem 7.2 (Follows from Peleg (1969)). *Every game on a (possibly infinite) graph has a Nash equilibrium.*

As already noted, we prove Theorem 7.2 by first using Lemma 2.5 to prove the existence of arbitrarily good approximate Nash equilibria, and then showing that the existence of such approximate Nash equilibria implies Theorem 7.2. For a given $\varepsilon > 0$, a mixed-strategy profile $(\sigma_i)_{i\in I}$ is an *ε-Nash equilibrium* if for every $i \in I$ and every possible deviating strategy $\sigma'_i \in \Sigma_i$, it holds that $u_i(\sigma_{N(i)}) \geq u_i(\sigma'_i, \sigma_{N(i)\setminus\{i\}}) - \varepsilon$.

Lemma 7.3. *For any $\varepsilon > 0$, every (possibly infinite) game on a graph has an ε-Nash equilibrium.*

[22]The converse does not hold for the analysis in that section, though: The proof there hinges on a full infinite sequence of approximations being encoded by a single model.

[23]Readers familiar with Peleg (1969) will note that even on graphs, Peleg's assumptions are weaker than those stated here. Our analysis can be generalized to cover such weaker assumptions.

Proof. Let $\varepsilon > 0$. For each player $i \in I$, it will be convenient to consider the space of profiles of mixed-strategies of players in $N(i)$ as a metric space with the ℓ^∞ metric. Note that this metric space is compact. As each player i has a continuous utility function whose domain is this compact metric space, players' utility functions are uniformly continuous by the Heine–Cantor theorem. Thus, there exists $\hat{\delta}_i > 0$ that assures that if two profiles of mixed strategies of players in $N(i)$ are less than $\hat{\delta}_i$ apart, then the utilities they yield to i differs by no more than $\varepsilon/2$.

For each player i, choose $\delta_i \triangleq \min\{\hat{\delta}_j \mid j \in N(i)\} > 0$. Recall that Σ_i denotes the space of player i's mixed strategies, and let $\Sigma_i^{\delta_i} \subset \Sigma_i$ be a finite set of strategies that includes all of i's pure strategies, and includes for any mixed strategy in Σ_i a strategy that is at most δ_i away from it; such a set exists by the compactness of Σ_i. We prove the lemma by proving that the given game admits an ε-Nash equilirbium in which each player i plays a strategy in $\Sigma_i^{\delta_i}$. We prove this using Lemma 2.5.

Definition of $\mathcal{P}$: Let $\mathcal{P}$ be all the subsets of I. A *solution* for $I' \in \mathcal{P}$ is a strategy profile for I that is a ε-Nash equilibrium in the induced game between all players in I' (where all other "players" play any arbitrary strategy), in which each player $i \in I$ plays a strategy in $\Sigma_i^{\delta_i}$.

Well describability: We introduce a variable $\texttt{plays}_{(i,\sigma_i)}$ for every player $i \in I$ and discretized strategy $\sigma_i \in \Sigma_i^{\delta_i}$. In what follows, for each $I' \in \mathcal{P}$ we define a set $\Phi_{I'}$ of formulae over these variables so that models of $\Phi_{I'}$ are in one-to-one correspondence with the (not-yet-proven-to-be-nonempty) set of solutions for I'. The correspondence is obtained by endowing the variable $\texttt{plays}_{(i,\sigma_i)}$ with the semantic interpretation "i plays the strategy σ_i." That is, it maps a model for Φ_I to the strategy profile such that for every $i \in I'$, we have that i plays the strategy σ_i if and only if the variable $\texttt{plays}_{(i,\sigma_i)}$ is True in that model. For every player i and every profile $\sigma_{N(i)\setminus\{i\}}$ of mixed-strategies for $N(i)\setminus\{i\}$, we define the set of ε-best responses of i:

$$\mathrm{BR}_i^\varepsilon(\sigma_{N(i)\setminus\{i\}}) \triangleq \Big\{\sigma_i \;\Big|\; u_i(\sigma_i, \sigma_{N(i)\setminus\{i\}}) \geq \max_{\sigma_i' \in \Sigma_i}\big\{u_i(\sigma_i', \sigma_{N(i)\setminus\{i\}})\big\} - \varepsilon\Big\}.$$

We define the set $\Phi_{I'}$ to consist of the following formulae:

1. for all $i \in I$, the (finite!) formula '$\bigvee_{\sigma\in\Sigma_i^{\delta_i}} \texttt{plays}_{(i,\sigma)}$', requiring that i plays some (discretized) strategy (this formula is finite because $\Sigma_i^{\delta_i}$ is);
2. for all $i \in I$ and all distinct $\sigma_i, \sigma_i' \in \Sigma_i^{\delta_i}$, the formula '$\texttt{plays}_{(i,\sigma_i)} \rightarrow \neg\texttt{plays}_{(i,\sigma_i')}$', requiring that the strategy that player i plays be unique;
3. for all $i \in I'$ and all profiles $\sigma = (\sigma_j)_{j\in N(i)\setminus\{i\}} \in \bigtimes_{j\in N(i)\setminus\{i\}} \Sigma_i^{\delta_i}$ of discretized

mixed strategies of $N(i) \setminus \{i\}$, the (finite!) formula

$$\Big(\bigwedge_{j\in N(i)\setminus\{i\}} \texttt{plays}_{(j,\sigma_j)}\Big) \to \Big(\bigvee_{\sigma_i\in\Sigma_i^{\delta_i}\cap \mathrm{BR}_i^{\varepsilon}(\sigma)} \texttt{plays}_{(i,\sigma_i)}\Big),$$

requiring that i ε-best-responds to the strategies played by the other players. By construction, $(\Phi_{I'})_{I'\in\mathcal{P}}$ is a well description of $\mathcal{P}$.

Finite-subset property: Let $\Phi' \subset \Phi_I$ be a finite subset. Since Φ' is finite, it "mentions" (through variables used) only finitely many players; denote the set of these players by $I' \subset I$. By definition, $\Phi' \subseteq \Phi_{I'}$. Consider the induced game on I' obtained by having each player $i \in I \setminus I'$ mechanically play some fixed strategy in $\Sigma_i^{\delta_i}$. By Theorem 7.1, this game has a Nash equilibrium. By choosing for each player $i \in I'$ a closest strategy in $\Sigma_i^{\delta_i}$ to the one she plays at this Nash equilibrium, each player's utility changes by at most $\varepsilon/2$ (by uniform continuity), and so does the utility attainable by best responding. Therefore, since we started with a Nash equilibrium, it is assured that each player is now playing an ε-best response, so the resulting strategy profile is a solution to I'. Therefore, I satisfies the finite-subset property. Thus, by Lemma 2.5, there exists an ε-Nash equilibrium in the grand game (among all player in I), as required. $\square$

Now, we can use Lemma 7.3 to prove Theorem 7.2 by way of a diagonalization argument.

Proof of Theorem 7.2. Since each player in the graph has finitely many neighbors, every connected component of the graph consists of at most countably many players. As it is enough to show the existence of a Nash equilibrium in each connected component separately (we use the Axiom of Choice here), let us focus on one connected component. By Lemma 7.3 there exists a sequence $(\sigma^n)_{n=1}^{\infty}$ of $\frac{1}{n}$-Nash equilibria in the game on this connected component. Since each of the at-most-countably-many coordinates of each element in this sequence lies in $[0, 1]$, we can choose a subsequence (a "diagonal subsequence") that converges in all coordinates; let σ^* denote the limit of that subsequence.

We claim that σ^* is a Nash equilibrium. To see this, note that for every $i \in I$ and $\sigma_i' \in \Sigma_i$, we have for the nth elements of the sequence that

$$u_i(\sigma^n_{N(i)}) \geq u_i(\sigma_i', \sigma^n_{N(i)\setminus\{i\}}) - \tfrac{1}{n}.$$

By the continuity of u_i, this means that for every $i \in I$ and $\sigma_i' \in \Sigma_i$, we have

$$u_i(\sigma^*_{N(i)}) \geq u_i(\sigma_i', \sigma^*_{N(i)\setminus\{i\}}),$$

so no player has a profitable deviation under the profile σ^*. Hence, σ^* is indeed a Nash equilibrium—and in particular, we see that a Nash equilibrium exists in the game, as desired. $\square$

7.2 Non-Applications

When using our framework, one faces an inherent tension. On the one hand, each formula in a well description must use finitely many variables. On the other hand, to be able to use finite results to establish the finite-subset property, each variable must be semantically related only to a finite set of elements in the economic problem. At first glance, this seems to preclude applications in which the desired solution has a parameter with an infinite domain, since requiring that the parameter take *some* value would require an infinite disjunction. Indeed, it is simpler to handle parameters with finite domains (which, as we have seen, naturally occur in many applications). Nevertheless, we have successfully applied Lemma 2.5 also to settings with infinite-domain parameters, such as utilities (Section 5), costs/prices (Appendix B), or probabilities (Section 7.1). Still, as the following examples demonstrate, in seemingly similar problems, this approach could not possibly work, since the infinite case has no solution.

Example 7.4 (Splitting the dollar)**.** A dollar must be split between a set I of agents. A solution is an efficient and envy-free division. When $|I| < \infty$ splitting the dollar equally is a solution. But the case $I = \mathbb{N}$ has no solution.

Example 7.5 (Higher number wins)**.** Two players can state a number in $S \subseteq \mathbb{R}$. The player whose stated number is higher wins a prize (which is shared in case of a tie). A solution is a pure-strategy Nash equilibrium. When $|S| < \infty$, a solution exists (each player states $\max S$), but the case $S = \mathbb{N}$ has no solution.

What are the limitations of our approach that prevent it from covering these two examples? Our approach for well-describing problems with infinite-domain parameters is to "encode" these parameters via a sequence of values, each from a finite domain. Specifically, we have encoded each of the abovementioned parameters using a sequence of increasingly fine discretizations.

Two features are critical for the success of this approach. First, that the desiderata on the encoded parameter can be imposed by individually finite formulae on the discretizations. For example, in Section 5, a utility function weakly rationalizes the data if and only if each of its discretizations weakly rationalize the data. This allowed us to represent a utility function that weakly rationalizes the data by a sequence of discretizations that each weakly rationalizes the data. The second critical feature is that not only each parameter value can be encoded by such a sequence of values (discretizations), but also each such sequence from any valid model encode a valid parameter value (i.e., in the parameter domain). In Section 5, since the sequence of discretized utilities is pointwise increasing and bounded, it has a finite limit. In our scaling of the rational inattention result of Caplin and Dean (2015) in Appendix B, since discretized costs are increasing, they have a limit, which is in $\mathbb{R}_{\geq 0} \cup \{\infty\}$—the domain of costs in that problem. For each of the above examples, there is no encoding that has both of these features.

In splitting the dollar, if, for example, we encode each player's allocation using discretizations, there is no set of individually finite formulae on the discretizations that holds if and only if the limit division is efficient. Hence, the first feature is missing. Any encoding that has this feature would lose the second feature.

In higher number wins, if, for example, we encode each player's number using discretizations, there is no set of individually finite formulae on the discretizations that holds if and only if the limit is finite. Hence, the second feature is missing. Any encoding that has this feature would lose the first feature.

8 Related Literature

In decision theory, both finite- and infinite-data models are important and common.[24] Reny (2015) showed how to unify these two approaches in the setting of Afriat (1967). In our view, our main contribution to this literature is in generalizing beyond any specific setting by providing a way to systematically unify these approaches. In Section 4, we proved that the result of Masatlioglu et al. (2012) scales to infinite datasets using essentially the same proof we used to reprove the classic result of Richter (1966) and Hansson (1968) that SARP suffices for rationalization by strict preferences. Like Masatlioglu et al., our infinite version of their theorem applies to *full* datasets. de Clippel and Rozen (2021) provide an analogous theorem for finite datasets that need not necessarily be full; our proof can be used to similarly scale their theorem to infinite datasets.

To our knowledge, we are the first to use Propositional Logic as a general tool for scaling results in economics. It is worth mentioning within this context, though, the work of Holzman (1984), who used Logical Compactness to relax topological conditions in Fishburn (1984).[25]

Our approach was stated using Propositional logic, but, in fact, Lemma 2.5 generalizes to well descriptions using First-Order Logic as well. Propositional logic is a special case of First-Order Logic. Importantly, it does not use quantifiers (i.e., $\forall$ and $\exists$). We chose to focus on this special case in order to simplify the exposition, since we were not able to identify any economic application in which the added

[24]For a recent example of a treatment of finite and infinite datasets, see Aguiar et al. (2020).

[25]Logical Compactness is frequently used to scale existence results in mathematics from finite settings to infinite ones (See, e.g., de Bruijn and Erdős, 1951 and Halmos and Vaughan, 1950).

generality would be beneficial.[26]

Other papers have used First-Order (rather than Propositional) Logic and non-standard analysis to unify, refine, and scale results in economic theory. Examples include Anderson (1978), Brown and Khan (1980), Anderson (1991), Khan (1993), Blume and Zame (1994), Halpern (2009), and Halpern and Moses (2016). Chambers et al. (2014) used Compactness in First-Order Logic to formalize the notion of the empirical content of a model. Like us, that paper studies applications to revealed preferences theory (see also Chambers et al., 2017), however it deals with different questions from us, and uses different techniques.

Hellman and Levy (2019) use (still different) tools from mathematical logic to prove conceptually related, yet incomparable, results: while our paper scales certain finite results to infinite settings, their paper scales certain countably infinite results to *un*countably infinite settings. Specifically, they give sufficient conditions to scale certain existence results that are known to hold whenever there are countably many possible *states of the world* into scenarios with uncountably many possible states of the world. Their results are incomparable to any of our results, and even to our existence-in-large-market results, first because they always assume that the number of agents is finite (an infinite number of agents, even with only two possible types for each, would already result in an uncountably infinite set of possible states of the world to begin with), and second, because they require that the theorems that they scale be already known to hold for the countably infinite, rather than only the finite, case.

We have been asked about the relation to various theorems in topology. Lemma 2.5 is stated in terms of logical propositions and its proof relies on Logical Compactness. Logical propositions can be translated into closed sets in an application-specific topological (product) space, in which setting Logical Compactness follows from Tychonoff's theorem on topological compactness. In other words, Lemma 2.5 can be proved using Tychonoff's theorem, and its statement can be translated to the language of topology. However, in our view, the resulting Lemma would be harder to directly formulate and the conditions would be harder to verify. And while topological compactness or the language of nets are stronger and more general approaches, in the domains we study, they often introduce technical issues that can render arguments incorrect in subtle ways (e.g., matchings may converge to an object that is not a matching). We therefore

[26] Well-describing economic problems using the full generality of First-Order Logic is Challenging. For example, fixing sets of objects (e.g., men and women) is not straightforward. In fact, by the (upward) Löwenheim–Skolem theorem, if a first-order theory has an infinite model (a model with an infinite domain) then it has a model of any larger cardinality, which implies that first-order theories cannot bound the cardinality of their infinite models. Hence, constants would have to play an important role in the well description.

view the methodological part of our contribution as introducing a unifying approach that is simple and intuitive to work with, and that does not require us to look for the "right" topological space or apply topological reasoning directly.[27]

9 Discussion

This paper provides a novel, principled approach for scaling economic theory results from finite models to infinite ones. We identify a sufficient condition for scaling a result: *A result can be scaled if it is well-describable with a description satisfying the finite-subset property.* The bulk of this paper is dedicated to demonstrating that many results in revealed-preference theory meet our condition, and therefore hold even with infinite datasets. We also demonstrate a game-theoretic application that focuses on a different "type" of scaling to infinity: allowing an infinite (rather than finite) number of players.

Our approach is not without limitation, and may fail where other approaches can succeed. That said, we have curated an array of applications showing that it has merit in decision theory and beyond in proving novel results, as well as consolidating and shortening proofs of previously known results, in a way that oftentimes sheds new light on them.

We view the main contribution of this paper to be a methodological one: a new, easy-to-use versatile tool for the economic theory toolbox. We hope that readers of this paper will be able to further leverage our approach.

References

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

V. Aguiar, P. Hjertstrand, and R. Serrano. A rationalization of the weak axiom of revealed preference. Working paper, Bravo Center for Economic Research, 2020.

R. M. Anderson. An elementary core equivalence theorem. *Econometrica*, pages 1483–1487, 1978.

R. M. Anderson. Non-standard analysis with applications to economics. In W. Hildenbrand and H. Sonnenschein, editors, *Handbook of Mathematical Economics*, volume 4, chapter 39, pages 2145–2208. Elsevier, 1991.

Y. Apartsin and Y. Kannai. Demand properties of concavifiable preferences. *Journal of Mathematical Economics*, 43(1):36–55, 2006.

R. J. Aumann. Values of markets with a continuum of traders. *Econometrica*, 43(4): 611–646, 1975.

[27]Once a proof is derived using Lemma 2.5, it is of course possible to then translate it to a topological statement and attempt to achieve greater generality, if/when such generality is of interest.

L. E. Blume and W. R. Zame. The algebraic geometry of perfect and sequential equilibrium. *Econometrica*, 62(4):783–794, 1994.

D. J. Brown and M. A. Khan. An extension of the Brown-Robinson equivalence theorem. *Applied Mathematics and Computation*, 6(2):167–175, 1980.

A. Caplin and M. Dean. Revealed preference, rational inattention, and costly information acquisition. *American Economic Review*, 105(7):2183–2203, 2015.

A. Caplin, M. Dean, and J. Leahy. Rationally inattentive behavior: Characterizing and generalizing Shannon entropy. Working paper, NBER, 2017.

M. D. Cattaneo, X. Ma, Y. Masatlioglu, and E. Suleymanov. A random attention model. *Journal of Political Economy*, 128(7):2796–2836, 2020.

C. P. Chambers and F. Echenique. Supermodularity and preferences. *Journal of Economic Theory*, 144:1004–1014, 2009.

C. P. Chambers and F. Echenique. *Revealed Preference Theory*. Cambridge University Press, 2016.

C. P. Chambers, F. Echenique, and E. Shmaya. The axiomatic structure of empirical content. *American Economic Review*, 104(8):2303–19, 2014.

C. P. Chambers, F. Echenique, and E. Shmaya. General revealed preference theory. *Theoretical Economics*, 12(2):493–511, 2017.

N. G. de Bruijn and P. Erdős. A colour problem for infinite graphs and a problem in the theory of relations. *Indigationes Mathematicae*, 13(5):371–373, 1951.

G. de Clippel and K. Rozen. Bounded rationality and limited data sets. *Theoretical Economics*, 16(2):359–380, 2021.

H. de Oliveira, T. Denti, M. Mihm, and K. Ozbek. Rationally inattentive preferences and hidden information costs. *Theoretical Economics*, 12(2):621–654, 2017.

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

H. Enderton. *A Mathematical Introduction to Logic*. Academic Press, 2nd edition, 2001.

E. Filiz-Ozbay and Y. Masatlioglu. Progressive random choice. Mimeo, 2020.

P. C. Fishburn. Comment on the Kannai-Peleg impossibility theorem for extending orders. *Journal of Economic Theory*, 32(1):176–179, 1984.

K. Gödel. Die vollständigkeit der axiome des logischen funktionenkalküls. *Monatshefte für Mathematik und Physik*, 37(1):349–360, 1930.

Y. A. Gonczarowksi and N. Nisan. *Mathematical Logic through Python*. Cambridge University Press, 2022.

Y. A. Gonczarowski, S. D. Kominers, and R. I. Shorrer. To infinity and beyond: A general framework for scaling economic theories. Working paper version of the present paper, arXiv:1906.10333, 2023.

L. J. Halbeisen. *Combinatorial Set Theory With a Gentle Introduction to Forcing.* Springer Monographs in Mathematics. Springer, 2nd edition, 2017.

P. R. Halmos and H. E. Vaughan. The marriage problem. *American Journal of Mathematics*, 72(1):214–215, 1950.

J. Y. Halpern. A nonstandard characterization of sequential equilibrium, perfect equilibrium, and proper equilibrium. *International Journal of Game Theory*, 38: 37–49, 2009. Erratum: *International Journal of Game Theory*, 46:591–594, 2017.

J. Y. Halpern and Y. Moses. Characterizing solution concepts in terms of common knowledge of rationality. *International Journal of Game Theory*, 46:457–473, 2016.

B. Hansson. Choice structures and preference relations. *Synthese*, 18(4):443–458, 1968.

Z. Hellman and Y. J. Levy. Measurable selection for purely atomic games. *Econometrica*, 87(2):593–629, 2019.

R. Holzman. An extension of fishburn's theorem on extending orders. *Journal of Economic Theory*, 32(1):192–196, 1984.

G. Hu, J. Li, J. Quah, and R. Tang. A theory of revealed indirect preference. Mimeo, 2021.

Y. Kannai. When is individual demand concavifiable? *Journal of Mathematical Economics*, 40(1-2):59–69, 2004.

M. Kearns. Graphical games. In N. Nisan, T. Roughgarden, É. Tardos, and V. V. Vazirani, editors, *Algorithmic Game Theory*, chapter 9, pages 159–178. Cambridge University Press, 2007.

M. A. Khan. The irony in/of economic theory. *MLN*, 108(4):759–803, 1993.

S. Lahiri. A simple proof of Suzumura's extension theorem for finite domains with applications. *Journal of Applied Mathematics and Decision Sciences*, 6(3):183–190, 2002.

A. Malcev. Untersuchungen aus dem gebiete der mathematischen logik. *Matematicheskii Sbornik*, 1(3):323–336, 1936.

M. Mandler. A quick proof of the order-extension principle. *American Mathematical Monthly*, 127(9):835–835, 2020.

A. Mas-Colell, M. D. Whinston, and J. R. Green. *Microeconomic theory.* Oxford University Press, 1995.

Y. Masatlioglu, D. Nakajima, and E. Y. Ozbay. Revealed attention. *American Economic Review*, 102(5):2183–2205, 2012.

A. Miralles and M. Pycia. Large vs. continuum assignment economies. Mimeo, 2015.

J. Nash. Non-cooperative games. *Annals of Mathematics*, 54(2):286–295, 1951.

E. A. Ok. *Real Analysis with Economic Applications.* Princeton University Press, 2007.

B. Peleg. Equilibrium points for games with infinitely many players. *Journal of the London Mathematical Society*, 44(1):292–294, 1969.

P. J. Reny. A simple proof of the nonconcavifiability of functions with linear not-all-parallel contour sets. *Journal of Mathematical Economics*, 49(6):506–508, 2013.

P. J. Reny. A characterization of rationalizable consumer behavior. *Econometrica*, 83 (1):175–192, 2015.

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

E. Szpilrajn. Sur l'extension de l'ordre partiel. *Fundamenta Mathematicae*, 1(16): 386–389, 1930.

A Proof of Theorem 4.2

Proof of Theorem 4.2. As with Theorem 3.2, the "only if" direction is immediate, so we prove the "if" direction using Lemma 2.5.

Definition of $\mathcal{P}$: Fixing X, let $\mathcal{P}$ be the set of all pairs (X', D) such that $X' \subseteq X$ and D is a full dataset over X' that satisfies WARP-LA. A *solution* for a pair $(X', D) \in \mathcal{P}$ is a pair of strict preference order over X' and attention filter that rationalize D.

Well describability: We introduce a variable ${}_a\mathtt{gt}_b$ for every pair of distinct $a, b \in X$ and a variable $\mathtt{attn}_{(S,T)}$ for each pair of finite sets S, T s.t. $\emptyset \neq T \subseteq S \subseteq X$. In what follows, for each $(X', D) \in \mathcal{P}$ we define a set $\Phi_{(X',D)}$ of formulae over these variables so that models of $\Phi_{(X',D)}$ are in one-to-one correspondence with the (not-yet-proven-to-be-nonempty) set of solutions for (X', D). The correspondence is obtained by endowing the variable ${}_a\mathtt{gt}_b$ with the semantic interpretation "a is preferred to b (when both are attention attracting)," and the variable $\mathtt{attn}_{(S,T)}$ with the semantic interpretation "T is the set of attention-attracting elements when the menu is S." That is, it maps a model for $\Phi_{(X',D)}$ to the preference $\succ$ such that for every distinct $a, b \in X'$, we have that $a \succ b$ if and only if the variable ${}_a\mathtt{gt}_b$ is True in that model and to the attention filter Γ such that for every S, T such that $\emptyset \neq T \subseteq S \subseteq X'$, we have that $\Gamma(S) = T$ if and only if the variable $\mathtt{attn}_{(S,T)}$ is True in that model. We define the set $\Phi_{(X',D)}$ to consist of the following formulae:

1. for all distinct $(S, a) \in D$, all $b \in S \setminus \{a\}$, and all $T \subseteq S$ s.t. $b \in T$, the formula '$\mathtt{attn}_{(S,T)} \rightarrow {}_a\mathtt{gt}_b$', requiring that the preferences and attention filter rationalize D;
2. for all distinct $a, b \in X'$, the formula '${}_a\mathtt{gt}_b \vee {}_b\mathtt{gt}_a$', requiring that the preferences be complete;
3. for all distinct $a, b \in X'$, the formula '$\neg({}_a\mathtt{gt}_b \wedge {}_b\mathtt{gt}_a)$', requiring that the preferences be antisymmetric;
4. for all distinct $a, b, c \in X'$, the formula '$({}_a\mathtt{gt}_b \wedge {}_b\mathtt{gt}_c) \rightarrow {}_a\mathtt{gt}_c$', requiring that the preferences be transitive;
5. for all menus $S \subseteq X'$, the (finite!) formula '$\vee_{\emptyset \neq T \subseteq S} \mathtt{attn}_{(S,T)}$', requiring that S have a set of attention-attracting elements that is a nonempty subset of S;
6. for all menus $S \subseteq X'$ and all distinct $T, T' \in 2^S \setminus \{\emptyset\}$, the formula '$\mathtt{attn}_{(S,T)} \rightarrow \neg\mathtt{attn}_{(S,T')}$', requiring that the set of attention-attracting elements from S be unique;
7. for all menus $S \subseteq X'$, all $T \in 2^S \setminus \{\emptyset\}$, and all $x \in S \setminus T$, the formula '$\mathtt{attn}_{(S,T)} \rightarrow \mathtt{attn}_{(S\setminus\{x\},T)}$', requiring that if x is not attention attracting from S, then S an $S \setminus \{x\}$ have the same attention-attracting set.

By construction, $(\Phi_{(X',D)})_{(X',D)\in\mathcal{P}}$ is a well description of $\mathcal{P}$.

Finite-subset property: Let $(\bar{X}, D) \in \mathcal{P}$. Let $\Phi' \subset \Phi_{(\bar{X},D)}$ be a finite subset. Since Φ' is finite, it "mentions" only finitely many elements of $\bar{X}$ (through variables used, whether by mentioning these elements directly or by mentioning menus that contain them); denote the set of these elements by $X' \subset \bar{X}$. Let $D' \triangleq \{(S,a) \in D \mid S \subseteq X'\}$. By definition, $\Phi' \subseteq \Phi_{(X',D')}$. Furthermore, D' is a full dataset and it satisfies WARP-LA since any sub-dataset of D satisfies WARP-LA. Hence, by Theorem 4.1, D' is rationalizable by some strict preference order over X' and attention filter. Therefore, $(\bar{X}, D)$ satisfies the finite-subset property. Thus, by Lemma 2.5, D is rationalizable by a strict preference order over $\bar{X}$ and an attention filter. □

B Rational Inattention

In this appendix, we demonstrate how our proof for Theorem 5.2 is quite a bit more flexible than one might imagine, by using essentially the same well description to scale quite a different rationalization result from finite to infinite datasets. For the most part, we follow the notation of Caplin and Dean (2015). Fix a finite set Ω of possible *states of the world* and a *prize space* X, as well as a *utility function* $u : X \to \mathbb{R}$. An *action* is a mapping from Ω to X. A *decision problem* is a finite set of actions. A *(state-dependent stochastic choice) dataset* is a collection D of decision problems along with functions $P_A : \Omega \to \Delta(A)$ for every $A \in D$, denoting the observed action distribution of a decision maker in each realized state of the world.

Fix a *prior belief* $\mu \in \Delta(\Omega)$ for the decision maker. An *information structure* is a distribution over a finite number of *posteriors* (distributions over Ω) for the decision maker whose average is μ.[28] The set of all information structures is Π. The utility from an action a when the posterior is γ is $g(a,\gamma) = \sum_{\omega\in\Omega} \gamma(\omega) u(a(\omega))$. The *gross payoff* from a decision problem A using an information structure π is $G(A,\pi) = \sum_{\gamma\in\operatorname{supp}\pi} \pi(\gamma) \max_{a\in A} g(a,\gamma)$. An *information cost function* is a mapping $K : \Pi \to \mathbb{R} \cup \{\infty\}$ where not all costs are infinite.

Given a dataset $(D, \{P_A\}_{A\in D})$, for every $a \in A$ by slight abuse of notation we write $P_A(a) = \sum_{\omega\in\Omega} \mu(\omega) P_A(a \mid \omega)$. For each $a \in A$ s.t. $P_A(a) > 0$, we write $P'_A(\cdot \mid a)$ for the *revealed posterior* associated with the action a in the decision problem A, defined by $P'_A(\omega \mid a) = \frac{P_A(a|\omega)\mu(\omega)}{P_A(a)}$ for every $\omega \in \Omega$.

A dataset $(D, \{P_A\}_{A\in D})$ is said to have *costly information acquisition representation* if there exists a cost function K such that for every problem $A \in D$:

1. For each $a \in A$ with $P_A(a) > 0$, it holds that $a \in \operatorname{argmax}_{b\in A} g(b, P'_A(\cdot \mid a))$.

[28] Caplin and Dean (2015) define an information structure as a mapping from Ω to distributions over posteriors that satisfiy Bayes' law with respect to the prior μ, which leads also to a different definition of rationalization; these definitions are known to be equivalent.

2. Letting π_A be the *revealed information structure* for A, i.e., the information structure such that for every $a \in A$ with $P_A(a) > 0$, the probability that π_A assigns to the posterior $P'_A(\cdot \mid a)$ is $P_A(a)$ (note that by definition, these probabilties sum up to 1, so this information structure is well defined), it holds that $\pi_A \in \operatorname{argmax}_{\pi\in\Pi}\big(G(A,\pi) - K(\pi)\big)$.

An information cost function K is *weakly monotone in information* if for every two information structures π, ϕ s.t. π is a garbling of ϕ, we have that $K(\phi) \geq K(\pi)$. An information cost function K is *mixture feasible* if for any two information structures π, ϕ and for every $\alpha \in (0,1)$, we have that $K\big(\alpha \circ \pi + (1-\alpha) \circ \phi\big) \leq \alpha K(\pi) + (1-\alpha)K(\phi)$, where $\alpha \circ \pi + (1-\alpha) \circ \phi$ is the *mixture distribution* assigning to each posterior a probability equal to the α-weighted average of the probabilities assigned to it by π and ϕ. Finally, an information cost function K is *normalized* if $K(\mathbb{1}_\mu) = 0$, where $\mathbb{1}_\mu$ is the information structure assigning probability 1 to the the prior μ being the posterior.

In a seminal result, Caplin and Dean (2015) showed that satisfying *No Improving Action Switches (NIAS)* and *No Improving Attention Cycles (NIAC)* is necessary and sufficient for having a costly information acquisition representation for a finite dataset—and furthermore that satisfying NIAS and NIAC is a sufficient condition for such a representation by a normalized, weakly monotone in information, and mixture feasible cost function. For our purposes, it suffices to note that whenever an infinite dataset satisfies NIAS and NIAC, so does any finite subset of it.

Theorem B.1 (Caplin and Dean, 2015)**.** *A finite dataset $(D, \{P_A\}_{A\in D})$ has a costly information acquisition representation if and only if it satisfies NIAS and NIAC. Moreover, when NIAS and NIAC hold there exists a costly information acquisition representation function for the dataset that is weakly monotone in information, mixture feasible, and normalized.*

Recently, Caplin et al. (2017) proved an infinite-dataset-size version of Theorem B.1 via a novel proof that diverges from Caplin and Dean's proof of the finite case. As we now show, this infinite version is also readily provable via Lemma 2.5, using Theorem B.1 for establishing the finite-subset property, with a very similar well description to the one we used to prove Reny's infinite-data version of Afriat's theorem. This is despite the differences between the two settings considered, and despite the fact that neither the original proofs of the finite versions nor the original proofs of the infinite versions of any of these quite different theorems share any common core technique.

Theorem B.2 (Caplin et al., 2017)**.** *A (possibly infinite) dataset $(D, \{P_A\}_{A\in D})$ has a costly information acquisition representation if and only if it satisfies NIAS* *and NIAC. Moreover, when NIAS and NIAC hold there exists a costly information acquisition representation function for the dataset that is weakly monotone in information, mixture feasible, and normalized.*

Proof of Theorem B.2. As noted, the "only if" direction is immediate, so we prove the "if" direction. Let $(D, \{P_A\}_{A \in D})$ be a dataset that satisfies NIAS and NIAC, and let π_A be the associated revealed information structure. Of the two conditions for having a costly information acquisition representation, the first condition does not involve K, and is in fact a condition on each decision problem in the data separately; the fact that it holds for each decision problem in the dataset separately follows from Theorem B.1 since each decision problem in the dataset satisfies NIAS and NIAC as a dataset by itself, and therefore has a costly information acquisition representation by itself as the only data point therefore satisfies the first condition for having a costly information acquisition representation.[29] The challenge with infinite data is therefore to find a cost function that satisfies the second condition for having a costly information acquisition representation, simultaneously for all infinitely many decision problems in the dataset. We do so using Lemma 2.5.

Definition of $\mathcal{P}$: Fixing X, Ω, and u, let $\mathcal{P}$ be the set of all datasets satisfying NIAS and NIAC. A *solution* for a dataset $(D, \{P_A\}_{A \in D}) \in \mathcal{P}$ is a costly information acquisition representation function for $(D, \{P_A\}_{A \in D})$ that is weakly monotone in information, mixture feasible, and normalized.

Well describability: We set $\varepsilon_n \triangleq 2^{-n}$ for every $n \in \mathbb{N}$. We introduce a variable $\texttt{cost}^n_{\pi,c}$ for every $n \in \mathbb{N}$, every information structure π, and every $c \in C_n \triangleq \{0, \varepsilon_n, 2 \cdot \varepsilon_n, \ldots, n\}$ (note that unlike in our proof of Theorem 5.2, this set goes up to n rather than only up to 1, so both the fineness of the discretization and the upper bound depend on n in this proof). In what follows, for each $(D, \{P_A\}_{A \in D}) \in \mathcal{P}$ we define a set $\Phi_{(D,\{P_A\}_{A \in D})}$ of formulae over these variables so that models of $\Phi_{(D,\{P_A\}_{A \in D})}$ are in one-to-one correspondence with the (not-yet-proven-to-be-nonempty) set of solutions for $(D, \{P_A\}_{A \in D})$. The correspondence is obtained by endowing the variable $\texttt{cost}^n_{\pi,c}$ with the semantic interpretation "$\lfloor K(\pi) \rfloor_{C_n} = c$ for the corresponding cost function K," where for every $n \in \mathbb{N}$ and every $x \in \mathbb{R} \cup \{\infty\}$ we denote by $\lfloor x \rfloor_{C_n}$ the rounding-down of x to the nearest number in C_n. (Note that in particular, $\lfloor x \rfloor_{C_n} = n$ for every $x > n$.) We define the set $\Phi_{(D,\{P_A\}_{A \in D})}$ to consist of the following formulae (which, except for the sixth formula-type, mirror those from our proof of Theorem 5.2):

1. for all $n \in \mathbb{N}$ and all information structures π, the (finite!) formula '$\bigvee_{c \in C_n} \texttt{cost}^n_{\pi,c}$', requiring that π have a rounded-down-to-C_n cost;

[29] In fact, only NIAS is required here, but we are intentionally stating the proof in a way that is agnostic to the details of NIAS and NIAC.

2. for all $n \in \mathbb{N}$, all information structures π, and all distinct $c, d \in C_n$, the formula '$\mathtt{cost}^n_{\pi,c} \to \neg\mathtt{cost}^n_{\pi,d}$', requiring that the rounded-down-to-C_n cost of π be unique;
3. for all $n \in \mathbb{N}$, all information structures π, and all $c \in C_n$, the following formula:
   - if $c < n$, the formula '$\mathtt{cost}^n_{\pi,c} \to \left(\mathtt{cost}^{n+1}_{\pi,c} \vee \mathtt{cost}^{n+1}_{\pi,c+\varepsilon_{n+1}}\right)$',
   - if $c = n$, the (finite) formula '$\mathtt{cost}^n_{\pi,c} \to \bigvee_{d=\{n,n+\varepsilon_{n+1},\ldots,n+1\}} \mathtt{cost}^{n+1}_{\pi,d}$',

   requiring that $\lfloor K(\pi)\rfloor_{C_n} = \lfloor\lfloor K(\pi)\rfloor_{C_{n+1}}\rfloor_{C_n}$;
4. For all $n \in \mathbb{N}$, all pairs of information structures π, ϕ s.t. π is a garbling of ϕ, and all $c \in C_n$, the (finite) formula '$\mathtt{cost}^n_{\pi,c} \to \bigvee_{d\in C_n:d\geq c} \mathtt{cost}^n_{\phi,d}$', requiring that the rounded-down-to-C_n cost function be weakly monotone in information;
5. for all $n \in \mathbb{N}$, all pairs of information strucures π, ϕ, all $\alpha \in (0,1)$, and all $c, d \in C_n$, the (finite) formula '$\left(\mathtt{cost}^n_{\pi,c} \wedge \mathtt{cost}^n_{\phi,d}\right) \to \bigvee_{c'\in C_n:c'\leq\alpha c+(1-\alpha)d+\varepsilon_n} \mathtt{cost}^n_{\alpha\circ\pi+(1-\alpha)\circ\phi,c'}$', requiring that the rounded-down-to-C_n cost function be mixture feasible, up to an error of at most ε_n;
6. for all $n \in \mathbb{N}$, the formula '$\mathtt{cost}^n_{\mathbb{1}_\mu,0}$', requiring that the rounded-down-to-C_n cost function be normalized (and, in particular, that not all costs are infinite);
7. for all decision problems $A \in D$, all $n \in \mathbb{N}$, all information structures ϕ, and all $c \in C_n \setminus \{n\}$, the (finite) formula '$\mathtt{cost}^n_{\phi,c} \to \bigvee_{d\in C_n:d\leq c-(G(A,\phi)-G(A,\pi_A))+\varepsilon_n} \mathtt{cost}^n_{\pi_A,d}$', requiring that choosing π_A maximizes gross payoff minus rounded-down-to-C_n costs, up to an error of at most ε_n.

We now argue that $(\Phi_{(D,\{P_A\}_{A\in D})})_{(D,\{P_A\}_{A\in D})\in\mathcal{P}}$ is a well description of $\mathcal{P}$. Let $(D, \{P_A\}_{A\in D}) \in \mathcal{P}$.

We first claim that every model that satisfies $\Phi_{(D,\{P_A\}_{A\in D})}$ corresponds to a solution for $(D, \{P_A\}_{A\in D})$. Fix a model for $\Phi_{(D,\{P_A\}_{A\in D})}$. For every information structure π and every $n \in \mathbb{N}$, let $c_n \in C_n$ be the value such that $\mathtt{cost}^n_{\pi,c_n}$ is True in the model (well defined by the first and second formula-types above), and define $K(\pi) = \lim_{n\to\infty} c_n$ (well defined, e.g., by the third formula-type above since c_n is monotone nondecreasing). The resulting cost function K is a limit of normalized functions that are weakly monotone in information (by the fourth and sixth formula-types above) and that up to an error that tends to 0, both (1) are mixture feasible (by the fifth formula-type above), and, (2) choosing π_A for each $A \in D$ maximizes gross payoff minus costs according to them (by the seventh formula-type above). Hence, K itself is normalized, weakly monotone in information, mixture feasible, and choosing π_A for each $A \in D$ maximizes gross payoff minus costs according to it (i.e., is a costly information acquisition representation function for $(D, \{P_A\}_{A \in D})$).

Second, if $(D, \{P_A\}_{A \in D})$ has a solution K, then using it we can construct a model for $\Phi_{(D,\{P_A\}_{A\in D})}$ (by setting each $\texttt{cost}^n_{\pi,c}$ to be True iff $c = \lfloor K(\pi) \rfloor_{C_n}$), and so $\Phi_{(D,\{P_A\}_{A\in D})}$ has a model. To sum up, $(\Phi_{(D,\{P_A\}_{A\in D})})_{(D,\{P_A\}_{A\in D})\in\mathcal{P}}$ is a well description of $\mathcal{P}$.

Finite-subset property: Let $(D, \{P_A\}_{A \in D}) \in \mathcal{P}$. Let $\Phi' \subset \Phi_{(D,\{P_A\}_{A\in D})}$ be a finite subset. Since Φ' is finite, there are only finitely many formulae of the above seventh type (the only formula type that depends on the dataset) in Φ'. For each of these formulae, select some decision problem in D that induces it, and let $D' \subset D$ be the set of these (at most $|\Phi'|$) problems. By definition, $\Phi' \subseteq \Phi_{(D',\{P_A\}_{A\in D'})}$. Furthermore, $(D', \{P_A\}_{A \in D'})$ satisfies NIAS and NIAC since any sub-dataset of $(D, \{P_A\}_{A \in D})$ satisfies NIAS and NAIC, and hence, by Theorem B.1, $(D', \{P_A\}_{A \in D'})$ has a solution. Therefore, $(D, \{P_A\}_{A \in D})$ satisfies the finite-subset property. Thus, by Lemma 2.5, $(D, \{P_A\}_{A \in D})$ has a solution. $\square$