

NESTED REPLICATOR DYNAMICS, NESTED LOGIT CHOICE, AND SIMILARITY-BASED LEARNING

PANAYOTIS MERTIKOPOULOS* AND WILLIAM H. SANDHOLM§

ABSTRACT. We consider a model of learning and evolution in games whose action sets are endowed with a partition-based similarity structure intended to capture exogenous similarities between strategies. In this model, revising agents have a higher probability of comparing their current strategy with other strategies that they deem similar, and they switch to the observed strategy with probability proportional to its payoff excess. Because of this implicit bias toward similar strategies, the resulting dynamics – which we call the *nested replicator dynamics* – do not satisfy any of the standard monotonicity postulates for imitative game dynamics; nonetheless, we show that they retain the main long-run rationality properties of the replicator dynamics, albeit at quantitatively different rates. We also show that the induced dynamics can be viewed as a stimulus-response model in the spirit of [Erev & Roth \(1998\)](#), with choice probabilities given by the nested logit choice rule of [Ben-Akiva \(1973\)](#) and [McFadden \(1978\)](#). This result generalizes an existing relation between the replicator dynamics and the exponential weights algorithm in online learning, and provides an additional layer of interpretation to our analysis and results.

1. INTRODUCTION

In the mass-action interpretation of evolutionary game theory, a nonatomic population of strategically-minded agents interact repeatedly via a set of simple myopic rules, and the governing dynamics arise from the aggregation of these individual interactions. Typically, these rules are described by a *revision protocol* – that is, a rule which determines how frequently agents playing a given strategy consider switching to a different one, and with what criteria. Thus, taken together, a population game and a revision protocol collectively define an aggregate dynamical system which uniquely prescribes the course of play starting from any initial population state.

A common feature of many revision protocols is the notion of *imitation*: the revising agent observes the behavior of a randomly selected individual, and then switches to the strategy of the observed agent with a probability that may depend on the revising agent’s incumbent payoff, the payoff of the observed strategy, or both. The archetypal example of this process is the *pairwise proportional imitation* (PPI) protocol of [Helbing \(1992\)](#), where the switching probability is proportional to the payoff excess of the observed strategy over the incumbent one, and the distribution of actions in the population evolves according to the replicator

*UNIV. GRENOBLE ALPES, CNRS, INRIA, GRENOBLE INP, LIG, 38000 GRENOBLE, FRANCE.

§DEPARTMENT OF ECONOMICS, UNIVERSITY OF WISCONSIN, 1180 OBSERVATORY DRIVE, MADISON WI 53706, USA.

E-mail addresses: panayotis.mertikopoulos@imag.fr, whs@ssc.wisc.edu.

2020 *Mathematics Subject Classification.* Primary; 91A22, 91A26; secondary 37N40, 68Q32.

Key words and phrases. Nested replicator dynamics; nested logit choice; nested pairwise proportional imitation; similarity-based learning; regularized learning.

The original idea for this paper was due to the second author, and the main results were obtained during a visit of the first author to UW Madison in 2017. Any mistakes or awkward turns of phrase are the sole responsibility of the first author.

dynamics of [Taylor & Jonker \(1978\)](#). Beyond this simple revision rule, different models of imitation have been considered by [Björnerstedt & Weibull \(1996\)](#), [Binmore & Samuelson \(1997\)](#), [Schlag \(1998\)](#), [Fudenberg & Imhof \(2006, 2008\)](#), [Mertikopoulos & Viossat \(2022\)](#), and many others; for a summary, see [Weibull \(1995\)](#), [Sandholm \(2010a\)](#), [Hadikhanloo et al. \(2022\)](#), and references therein.

Invariably, in all these revision rules, the revising agent inspects the strategy of “the first person they meet in the street” ([Weibull, 1995](#)). This premise is certainly reasonable when there are no prior beliefs or selection biases in the population; however, when the players’ strategies share a set of exogenous similarities, an agent may be biased toward sampling similar strategies rather than dissimilar ones. [For example, when commuting from home to work, an agent may prefer to compare routes or departure times for the same mode of transport instead of a different, less familiar one.] We are thus led to the following question:

*How are the dynamics affected if players are more likely to observe
and compare strategies that are more similar to their current ones?*

Our contributions. To address this question, we examine a simple model of evolution and learning in games whose action sets are endowed with a partition-based *similarity structure*, i.e., a hierarchy of nested partitions – or *similarity classes* – that capture the degree of exogenous similarity between distinct strategies. In our model, the revising agent has a higher probability of comparing payoffs with agents playing similar strategies, and the greater the degree of similarity between strategies, the greater the probability of such a comparison being made. Then, after observing their opponent’s payoff, the agent considers imitating the observed strategy only if its payoff is higher than their current one, and they consummate the switch with probability proportional to the difference in payoffs.

When there are no exogenous similarities between strategies, this model reduces to the pairwise proportional imitation protocol of [Helbing \(1992\)](#) mentioned above, and the population follows the replicator dynamics of [Taylor & Jonker \(1978\)](#). In our approach, the additional affinity toward similar, more familiar strategies introduces an implicit selection bias, which in turn leads to quite different aggregate dynamics. We call the resulting law of evolution the *nested replicator dynamics* (NRD), in reference to the nested similarity structure that drives the players’ payoff comparisons.¹

To have a baseline in mind, the standard replicator dynamics exhibit a wide range of rationality properties, including the elimination of strictly dominated strategies, the local stability of evolutionarily stable states, global convergence in potential and strictly contractive games, etc.² A subset of these properties extend to a more general class of dynamics exhibiting monotone percentage growth rates, the canonical monotonicity condition for dynamics based on imitation.³ Nevertheless, because of the increased affinity toward similar strategies, the nested replicator dynamics fail even the most basic versions of monotonicity: unless the implicit bias toward similar strategies is very small, a strategy that performs well relative to the full collection of strategies but poorly compared to similar strategies will tend to become less common in the population. It is therefore unclear whether the nested replicator dynamics possess *any* appealing rationality properties whatsoever.

Despite this handicap, we show that the nested replicator dynamics provide remarkably strong support for traditional solution concepts. More precisely, our first result ([Theorem 1](#)) shows that they retain all the mainstay rationality properties of the replicator dynamics,

¹[Mertikopoulos & Sandholm \(2018\)](#) derived a version of these dynamics as a special case of a “Hessian Riemannian system”, and deferred the dynamics’ microfoundations and analysis to the current paper. This link with Riemannian game dynamics can be seen as an additional layer of interpretation.

²For a detailed discussion, see [Hofbauer & Sigmund \(1988\)](#), [Weibull \(1995\)](#) and [Sandholm \(2010a, 2015\)](#).

³For early appearances of this, see [Nachbar \(1990\)](#); [Friedman \(1991\)](#) and [Samuelson & Zhang \(1992\)](#).

including in particular that (i) strictly dominated strategies become extinct in the long run; (ii) if the dynamics converge, their limit is a Nash equilibrium; (iii) dynamically stable states are Nash equilibria; (iv) strict Nash equilibria and, more generally, evolutionarily stable states (ESSs) are dynamically stable and attracting (with global ESSs being globally attracting); and (v) the dynamics converge to Nash equilibrium in potential and strictly contractive games. However, from a quantitative viewpoint, there are visible differences in the rate at which these properties emerge, e.g., as a function of whether a strategy is dominated by a similar or a dissimilar strategy.

In addition to the above, we also establish a fundamental connection between the nested replicator dynamics and stimulus-response learning in the spirit of [Erev & Roth \(1998\)](#). The precise setting of this connection is a continuous-time model of learning through reinforcement in which agents track the cumulative payoffs of their strategies and select strategies by feeding these cumulative scores into a probabilistic choice rule. It is known from the work of [Rustichini \(1999\)](#) and [Hofbauer et al. \(2009\)](#) that if players choose mixed strategies using the logit choice rule, the evolution of these strategies follows the replicator dynamics. In [Section 5](#), we derive a considerable generalization of this result: if players employ the *nested logit choice* rule ([Ben-Akiva, 1973](#); [McFadden, 1978](#)) their mixed strategies follow the *nested replicator dynamics* ([Theorem 2](#)).

This relation is particularly surprising in light of the fundamental differences between the two nesting constructions. This is partially explained by our results in [Section 6](#), where we describe a variant model of similarity-based learning, and we provide a nonrecursive representation of the nested logit choice rule as a “regularized best-response” model. This last representation shows that the nested replicator dynamics can also be seen as a special case of the widely used “follow the regularized leader” (FTRL) family of online learning methods ([Shalev-Shwartz, 2011](#)), providing in this way a triple equivalence between the revisionist, stimulus-response, and regularized learning viewpoints.

Related work on similarity-based reasoning. The idea that similarities play an important role for decision-making goes back at least to [Luce \(1956\)](#), with [Debreu \(1960\)](#) pointing out that realistic models of probabilistic choice should allow for violations of the principle of “irrelevance of added alternatives” ([Luce, 1959](#)). This issue is illustrated clearly by a canonical example from transportation due to [McFadden \(1974b\)](#), where two similar options (a blue bus and a red bus) are each chosen less frequently than a dissimilar one (a car), even when all options are otherwise equally desirable in terms of utilities. In this regard, the need to account for similarities when taking decisions has had a lasting impact in lasting impact in theoretical and econometric models of discrete choice ([McFadden, 1981, 2001](#); [Ben-Akiva & Lerman, 1985](#); [Anderson et al., 1992](#)).

In a learning context, [Gilboa & Schmeidler \(1995\)](#) considered a model where agents reason by drawing analogies to similar situations encountered in the past, and they provided a concrete axiomatization of decision rules that output the “best” action based on its past performance in similar cases. More recently, [Steiner & Stewart \(2008\)](#) examined an analogous setting where players continually encounter new strategic, normal-form games and extrapolate beliefs from similar past situations, ultimately generating contagion of actions across different games through a “transfer learning” process. This idea was taken further by [Mengel \(2012\)](#) who introduced a partition of all games encountered into categories, and considered a process of simultaneous learning of actions and partitions. A crucial point where our conclusions differ with these works is that the dynamics of [Mengel \(2012\)](#) *destabilize* strict equilibria, whereas strict equilibria *are always stable* under the nested replicator dynamics.

One reason for this discrepancy is that the similarities mentioned above concern events and situations encountered in the past, not the players’ actions per se – in this sense, similarities

are *endogenous*, not *exogenous*. In the context of extensive-form games, Jehiel (2005) – and, more recently, Jehiel & Samet (2005, 2007) – examined a dynamic behavioral model where agents bundle together decision nodes into distinct, static similarity classes, and agents only try to learn based on the average behavior in each class. While the model of Jehiel & Samet (2005, 2007) is related in spirit to our approach (in both cases, the notion of similarity is determined exogenously and it does not fluctuate over time), their focus is on partitions of sequences of moves in extensive-form games, which have no analogue in the partition of one-shot actions in a population game.

Finally, in a recent unpublished manuscript, Börgers & Mailath (2020) introduced a model of learning in finite games with exogenous similarities in the signals received by the players.⁴ These similarities need not be transitive or symmetric (so they do not necessarily comprise a partition-based structure), leading to the novel solution concept of a *similarity equilibrium*.⁵ Börgers & Mailath (2020) subsequently showed that similarity-based learning only converges to similarity equilibria – i.e., if the players’ learning process converges, the limit is a similarity equilibrium. This property is reminiscent of the internal consistency of the nested replicator dynamics – i.e., if the nested replicator dynamics converge, they converge to a Nash equilibrium – but the setting of Börgers & Mailath (2020) is otherwise quite different to our own, so no further analogies can be drawn.

2. PRELIMINARIES

2.1. Population games. Throughout this paper, we focus on single-population games, i.e., the set of *players* is modeled by the unit interval $\mathcal{I} = [0, 1]$ and each agent selects (in a measurable way) an *action* – or *pure strategy* – from some finite set $\mathcal{A} = \{\alpha_i : i = 1, \dots, m\}$.⁶ The players’ payoffs are then determined by their choice of action and the population shares of each action (so any measure-zero set of players has no impact on the game’s payoffs).

In more detail, if the mass of agents playing $\alpha \in \mathcal{A}$ is $x_\alpha \in [0, 1]$, the *state of the population* is defined as the overall distribution of actions $x = (x_\alpha)_{\alpha \in \mathcal{A}}$, viewed here as an element of the m -simplex

$$\mathcal{X} \equiv \Delta(\mathcal{A}) = \{x \in \mathbb{R}_+^m : \sum_{\alpha \in \mathcal{A}} x_\alpha = 1\}. \quad (1)$$

For posterity, given a subset \mathcal{S} of \mathcal{A} , we will also write

$$x_{\mathcal{S}} = \sum_{\alpha \in \mathcal{S}} x_\alpha \quad (2)$$

for the mass of agents playing actions drawn from \mathcal{S} , and, analogously, we will write

$$x_{\mathcal{S}'|\mathcal{S}} = \frac{x_{\mathcal{S} \cap \mathcal{S}'}}{x_{\mathcal{S}}} = \frac{\sum_{\alpha \in \mathcal{S} \cap \mathcal{S}'} x_\alpha}{\sum_{\alpha \in \mathcal{S}} x_\alpha} \quad (3)$$

for the relative share of agents drawing an action from $\mathcal{S}' \subseteq \mathcal{A}$ within the sub-population of agents playing an action in \mathcal{S} (obviously, $x_{\mathcal{S}'|\mathcal{S}} = 0$ if $\mathcal{S} \cap \mathcal{S}' = \emptyset$).⁷

The nonatomic nature of the game is captured by the fact that the players’ rewards only depend on the state of the population and not the individual choices of each player. More precisely, the players’ rewards are determined by an ensemble of *payoff functions* $v_\alpha : \mathcal{X} \rightarrow \mathbb{R}$, $\alpha \in \mathcal{A}$, with $v_\alpha(x)$ representing the payoff to α -strategists when the population’s state is

⁴We thank Tilman Börgers for bringing this work to our attention; our paper’s title is partially due to this work.

⁵When the similarities are partition-based, similarity equilibria boil down to correlated equilibria.

⁶Our analysis extends to multi-population settings without much effort, but we do not treat this case to avoid overburdening the notation and the presentation.

⁷For simplicity, we will make no distinction between $x_{\{\alpha\}}$ and x_α for any individual action $\alpha \in \mathcal{A}$. We also assume above that x has full support, i.e., $x_\alpha > 0$ for all $\alpha \in \mathcal{A}$.

x . Collectively, we will write $v(x) = (v_\alpha(x))_{\alpha \in \mathcal{A}}$ for the population's *payoff vector* at state $x \in \mathcal{X}$, so the associated population mean payoff is given by

$$u(x) = \langle v(x), x \rangle = \sum_{\alpha \in \mathcal{A}} x_\alpha v_\alpha(x), \quad (4)$$

with $\langle \pi, x \rangle \equiv \sum_{\alpha \in \mathcal{A}} x_\alpha \pi_\alpha$ denoting the canonical pairing between π and x (both viewed here as elements of \mathbb{R}^m). Putting all this together, a *population game* is a tuple $\mathcal{G} \equiv \mathcal{G}(\mathcal{A}, v)$ with \mathcal{A} and v defined as above.

Example 1 (Symmetric random matching). One of the most widely studied classes of population games is given by *random matching* in finite games (Weibull, 1995; Sandholm, 2010a). In this setting, two players are drawn uniformly at random from the population and, subsequently, they are matched to play a symmetric two-player game with payoff matrix $M \in \mathbb{R}^{m \times m}$. Thus, averaging over the population, the mean payoff to α -strategists at state $x \in \mathcal{X}$ is $v_\alpha(x) = \sum_{\beta \in \mathcal{A}} M_{\alpha\beta} x_\beta$ and the game's payoff field is $v(x) = Mx$. §

2.2. Solution concepts. In our analysis, we will consider an array of different solution concepts and notions for nonatomic games. For completeness, we recall some basic definitions below:

Definition 1. Consider a population game $\mathcal{G} \equiv \mathcal{G}(\mathcal{A}, v)$. Then:

- (1) A pure strategy $\alpha \in \mathcal{A}$ is *strictly dominated* by $\beta \in \mathcal{A}$ if

$$v_\alpha(x) < v_\beta(x) \quad \text{for all } x \in \mathcal{X}. \quad (5)$$

- (2) A population state $x^* \in \mathcal{X}$ is a *Nash equilibrium* (NE) of \mathcal{G} if

$$v_\alpha(x^*) \geq v_\beta(x^*) \quad \text{for all } \alpha, \beta \in \mathcal{A} \text{ such that } x_\alpha^* > 0 \quad (6)$$

or, equivalently, if

$$\langle v(x^*), x - x^* \rangle \leq 0 \quad \text{for all } x \in \mathcal{X}. \quad (7)$$

Finally, x^* is a *restricted equilibrium* of \mathcal{G} if $v_\alpha(x^*) = v_\beta(x^*)$ for all $\alpha, \beta \in \text{supp}(x^*)$.

- (3) A population state $x^* \in \mathcal{X}$ is an *evolutionarily stable state* (ESS) of \mathcal{G} if

$$\langle v(x), x - x^* \rangle < 0 \quad \text{for all } x \neq x^* \text{ close to } x^*. \quad (8)$$

In particular, x^* is a *globally evolutionarily stable state* (GESS) if (8) holds for all $x \in \mathcal{X} \setminus \{x^*\}$.

- (4) \mathcal{G} is a *potential game* if $v(x) = \nabla F(x)$ for some *potential function* $F: \mathcal{X} \rightarrow \mathbb{R}$.

- (5) \mathcal{G} is *strictly monotone* – or *strictly contractive* – if

$$\langle v(x') - v(x), x' - x \rangle \leq 0 \quad \text{for all } x, x' \in \mathcal{X}, \quad (9)$$

with equality if and only if $x = x'$.⁸

Some well-known facts that will come in handy later are as follows:

- If x^* is evolutionarily stable, it is *a fortiori* a Nash equilibrium (Hofbauer et al., 1979).
- If \mathcal{G} is potential, any local maximizer of F is a Nash equilibrium (Sandholm, 2001).
- If a game admits a (strictly) concave potential, it is (strictly) monotone; also, a strictly monotone game admits a unique Nash equilibrium which is a GESS (Sandholm, 2010a).

For an in-depth discussion, see Weibull (1995), Sandholm (2010a) and references therein.

⁸The terminology “contractive” follows Sandholm (2015); by contrast, Hofbauer & Sandholm (2009) and Sorin & Wan (2016) respectively use the term *stable* and *dissipative* for essentially the same concept. Our choice of terminology reflects the long history of (9) in convex analysis and optimization, where it is known as “*monotonicity*”, cf. Hadikhani et al. (2022), Duvocelle et al. (2023), and references therein. The first author hopes that the second author would have found this compromise acceptable.

2.3. Revision protocols and evolutionary dynamics. Evolutionary game dynamics are commonly derived from explicit microfoundations describing the agents' efforts to increase their individual payoffs (Weibull, 1995; Benaïm & Weibull, 2003; Sandholm, 2010a,b). In this context, each agent occasionally receives opportunities to switch actions – say, based on the rings of a Poisson alarm clock – and, at such moments, they choose a new action by applying a *revision protocol*.

Typically, a revision protocol is defined by specifying the *conditional switch rates* $\varrho_{\alpha\beta}(\pi, x)$ at which an α -strategist might consider switching to strategy β when the population is at state $x \in \mathcal{X}$ and the payoff vector is $\pi \in \mathbb{R}^m$. In this way, for any given population game $\mathcal{G} \equiv \mathcal{G}(\mathcal{A}, v)$, a revision protocol ϱ induces the *mean dynamics*

$$\dot{x}_\alpha = \sum_{\beta \in \mathcal{A}} x_\beta \varrho_{\beta\alpha}(v(x), x) - x_\alpha \sum_{\beta \in \mathcal{A}} \varrho_{\alpha\beta}(v(x), x). \quad (\text{MD})$$

The dynamics (MD) specify the absolute rate of change in the use of each action $\alpha \in \mathcal{A}$ as the difference between inflows *to* α from other actions and outflows *from* α to other actions. To motivate the analysis to come, we show how this approach provides microfoundations for the *replicator dynamics*, the most widely studied dynamics in evolutionary game theory.

Example (Pairwise proportional imitation). Following Helbing (1992), consider the *pairwise proportional imitation* protocol

$$\varrho_{\alpha\beta}(\pi, x) = x_\beta [\pi_\beta - \pi_\alpha]_+ \quad (\text{PPI})$$

where $[x]_+ := \max\{0, x\}$ denotes the positive part of $x \in \mathbb{R}$. Under (PPI), a revising α -strategist meets another individual – a *mentor* or *comparator* – drawn from the population uniformly at random; thus, if the population is at state $x \in \mathcal{X}$, a β -strategist is encountered with probability x_β . If the payoff of the revising agent is lower than the comparator's, the agent switches to the comparator strategy with probability proportional to the payoff excess $[\pi_\beta - \pi_\alpha]_+$; otherwise, the revising agent sticks to their current choice. Accordingly, the *conditional imitation rate* of β -strategists by α -strategists under (PPI) is

$$r_{\alpha\beta}(\pi, x) = [\pi_\beta - \pi_\alpha]_+ \quad (10)$$

so $\varrho_{\alpha\beta} = x_\beta r_{\alpha\beta}$. Then, substituting protocol (PPI) into (MD) and rearranging yields the *replicator dynamics* of Taylor & Jonker (1978), viz.

$$\dot{x}_\alpha = x_\alpha [v_\alpha(x) - u(x)] \quad (\text{RD})$$

with $u(x) = \sum_{\alpha \in \mathcal{A}} x_\alpha v_\alpha(x)$ denoting the mean population payoff under x . This derivation plays a central role in our analysis, so we will return to it several times in the sequel. §

3. THE MODEL: IMITATION DRIVEN BY SIMILARITIES

We now turn to our paper's main objective, that is, the study of how similarities between strategies could impact the players' aggregate behavior over time. Our model for this has two main components: (i) a *similarity structure* intended to capture varying degrees of similarity between strategies; and (ii) the *nested pairwise proportional imitation* (NPPI) protocol, which provides a similarity-driven variant of the PPI protocol of Helbing (1992). Taken together, these two components define the *nested replicator dynamics* (NRD), which we introduce later in this section and analyze in detail in Section 4.

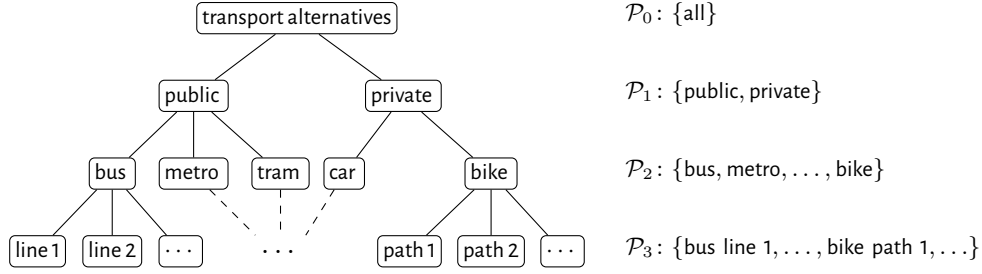


Figure 1: Grouping of alternative modes of transport by similarity.

3.1. Similarity structures. To set the stage for the sequel – and because the notation involved can become somewhat cumbersome at times – it will help to keep in mind as a running example a population of commuters who consider different alternatives to go from home to work each morning. At the highest level, these alternatives could be grouped into public and private transport modes; subsequently, different alternatives could be divided up into smaller subgroups based on the individual transport mode chosen (bus, car, metro, ...), the choice of route and/or departure time, etc. The result is a tree-like grouping of alternatives into nested clusters and subclusters, which we depict in Fig. 1 above.

To put all this on a formal footing, we will assume that the players’ set of strategies comes equipped with a hierarchy of increasingly finer partitions as follows:

Definition 2. A *similarity structure* – or *hierarchy* – on \mathcal{A} is a tower of nested *similarity partitions* \mathcal{P}_ℓ , $\ell = 0, \dots, L$, of \mathcal{A} such that

$$\{\mathcal{A}\} =: \mathcal{P}_0 \succ \mathcal{P}_1 \succ \dots \succ \mathcal{P}_L := \{\{\alpha\} : \alpha \in \mathcal{A}\} \quad (11)$$

where the notation “ $\mathcal{P} \succ \mathcal{Q}$ ” means that the partition \mathcal{Q} is *finer* than \mathcal{P} , i.e., every element of \mathcal{Q} is a subset of some element of \mathcal{P} .⁹

In this definition, each partition \mathcal{P}_ℓ , $\ell = 1, \dots, L$, captures successively finer traits of the elements of \mathcal{A} : for instance, in the commuting example of Fig. 1, the coarsest non-trivial partition \mathcal{P}_1 would correspond to the distinction between public and private transport modes; the second partition $\mathcal{P}_2 \preccurlyeq \mathcal{P}_1$ to different individual modes (bus, metro, car, ...); the third partition $\mathcal{P}_3 \preccurlyeq \mathcal{P}_2$ to different routes for each mode; etc. In this interpretation, the index $\ell = 1, \dots, L$ indicates the degree of similarity between two strategies: two strategies that cannot be resolved by \mathcal{P}_ℓ are said to be *similar at level ℓ* ,¹⁰ so the strategies that are similar at level $L - 1$ are the *most similar* (because they share all attributes of \mathcal{P}_{L-1}), whereas the strategies that are not similar even at level 1 are the *most dissimilar* ones (since they do not share any common attributes whatsoever). Thus, in our running example, two different bus lines have 2 degrees of similarity, but a bike path and a bus line are not similar at any level.

For concreteness, each constituent set \mathcal{S} of a partition \mathcal{P}_ℓ will be referred to as a *similarity class of level ℓ* . A similarity structure on \mathcal{A} may thus be represented as a disjoint union of all class/level pair (\mathcal{S}, ℓ) for $\mathcal{S} \in \mathcal{P}_\ell$, viz.

$$\mathcal{P}_\bullet := \coprod_{\ell=0}^L \mathcal{P}_\ell \equiv \bigcup_{\ell=0}^L \{(\mathcal{S}, \ell) : \mathcal{S} \in \mathcal{P}_\ell\}. \quad (12)$$

⁹The trivial partition $\mathcal{P}_0 = \{\mathcal{A}\}$ does not carry any information in itself, but it has been included for completeness and notational convenience later on.

¹⁰By “resolve” we mean here that the two strategies in question belong to different elements of \mathcal{P}_ℓ . By construction, any two distinct strategies can be resolved by \mathcal{P}_L , but no strategies can be resolved by \mathcal{P}_0 .

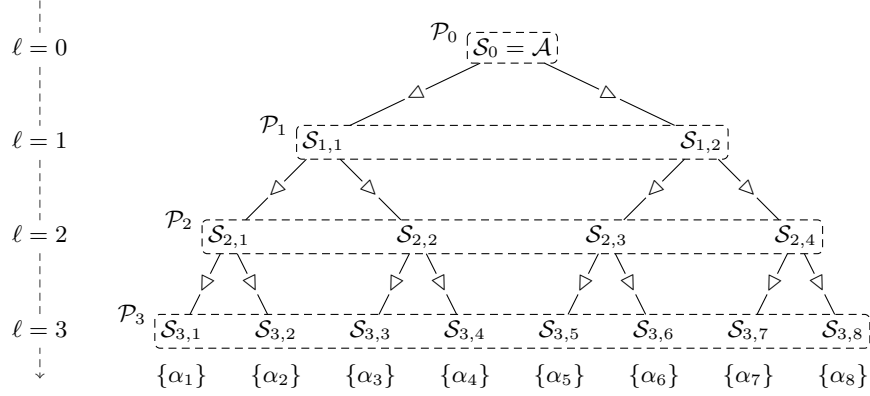


Figure 2: Graphical representation of a 3-tier similarity structure as a rooted tree.

To simplify notation, if there is no danger of confusion, we will not differentiate between the class \mathcal{S} and the class/level pair (\mathcal{S}, ℓ) ; by contrast, when we need to unambiguously keep track of the level index ℓ of \mathcal{S} , we will write $\ell = \text{lvl}(\mathcal{S})$.

Moving forward, if a class $\mathcal{S} \in \mathcal{P}_\ell$ contains the class $\mathcal{S}' \in \mathcal{P}_k$ for some $k > \ell$, we will say that \mathcal{S}' is a *descendant* of \mathcal{S} and we will write “ $\mathcal{S}' \prec \mathcal{S}$ ” (resp. “ $\mathcal{S} \succ \mathcal{S}'$ ” when \mathcal{S} is an *ancestor* of \mathcal{S}'). As a special case of this relation, if $\mathcal{S}' \prec \mathcal{S}$ and $k = \ell + 1$, we will say that \mathcal{S}' is a *child* of \mathcal{S} and we will write “ $\mathcal{S}' \triangleleft \mathcal{S}$ ” (resp. \mathcal{S} is a *parent* of \mathcal{S}' when $\mathcal{S} \triangleright \mathcal{S}'$). More generally, when we wish to focus on descendants sharing a certain attribute, we will write “ $\mathcal{S}' \preceq_\ell \mathcal{S}$ ” as shorthand for the predicate “ $\mathcal{S}' \preceq \mathcal{S}$ and $\text{lvl}(\mathcal{S}') = \ell$ ”. In a similar vein, given an arbitrary class $\mathcal{S} \in \mathcal{P}_\bullet$, we will write $[\mathcal{S}]_\ell$ for the ℓ -th type ancestor of \mathcal{S} , i.e., the (necessarily unique) similarity class $\mathcal{S}_\ell \in \mathcal{P}_\ell$ containing \mathcal{S} (obviously, we assume here that $\text{lvl}(\mathcal{S}) \geq \ell$). Finally, we will say that two classes $\mathcal{S}, \mathcal{S}' \in \mathcal{A}$ are *similar with respect to \mathcal{P}_ℓ* if and only if $[\mathcal{S}]_\ell = [\mathcal{S}']_\ell$, in which case we will write $\mathcal{S} \stackrel{\ell}{\sim} \mathcal{S}'$.¹¹

Building on the above, a similarity structure on \mathcal{A} can be represented graphically as a rooted directed tree – or *arborescence* – by connecting two classes $\mathcal{S}, \mathcal{S}' \in \mathcal{P}_\bullet$ with a directed edge $\mathcal{S} \rightarrow \mathcal{S}'$ whenever $\mathcal{S} \triangleright \mathcal{S}'$ (cf. Fig. 2 above). By construction, the root of this tree is \mathcal{A} itself, and the unique directed path $\mathcal{A} \equiv \mathcal{S}_0 \triangleright \mathcal{S}_1 \triangleright \dots \triangleright \mathcal{S}_\ell \equiv \mathcal{S}$ from \mathcal{A} to any class $\mathcal{S} \in \mathcal{P}_\bullet$ will be referred to as the *lineage* of \mathcal{S} . For notational simplicity, we will not distinguish between \mathcal{P}_\bullet and its graphical representation, and we will use the two interchangeably.

Example. Going back to the commuting example of Fig. 1, the class “bus” is a child of the parent class “public” (formally, $\text{bus} \triangleleft \text{public}$), and it is a sibling / similar to the classes “metro” and “tram”, both of which are also children of the class “public” (so $\text{bus} \stackrel{2}{\sim} \text{metro} \stackrel{2}{\sim} \text{tram}$ in the hierarchy depicted in Fig. 1). Accordingly, the lineage of the class “bus” is the two-level parent/child chain $\text{all} \triangleright \text{public} \triangleright \text{bus}$, where the class “all” bundles together all transportation modes in a single (trivial) supercluster. §

3.2. Nested pairwise proportional imitation and the nested replicator dynamics. Building on the PPI protocol described in the previous section, we introduce below a more general family of revision protocols under which agents give additional consideration to switches to “similar”, more familiar actions. For example, a commuter may be more likely to contemplate switching to a different instance of their current transport mode than a different mode altogether (e.g., taking a different combination of metro lines rather than switching to a bus,

¹¹The notations $[\alpha]$ and $\alpha \sim \beta$ will be reserved for elements of \mathcal{A} (i.e., when $\ell = L - 1$).

cf. Ben-Akiva, 1973; Ben-Akiva & Lerman, 1985; McFadden, 1981, 2001). Likewise, when deciding to replace a computer or other technology with network externalities, a consumer may be more likely to consider models manufactured by the maker of their current machine than those of other manufacturers, particularly in contexts with switching costs (Katz & Shapiro, 1994; Farrell & Klemperer, 2007); etc.

Formally, given a similarity structure $\mathcal{P}_0 \succ \mathcal{P}_1 \succ \dots \succ \mathcal{P}_L$ of \mathcal{A} as above, we will consider a multi-stage revision protocol that unfolds as follows:

- (1) When given an opportunity to switch strategies, a revising agent playing $\alpha \in \mathcal{A}$ first decides which strategies to consider as candidates for imitation. To account for the increased familiarity of more similar strategies, the pool of candidates from α is

$$\text{cand}(\alpha) = \begin{cases} [\alpha]_0 & \text{with probability } \lambda_0, \\ [\alpha]_1 & \text{with probability } \lambda_1, \\ \dots & \\ [\alpha]_{L-1} & \text{with probability } \lambda_{L-1}, \end{cases} \quad (13)$$

where the “intra-level” sampling probabilities $\lambda_k \geq 0$ satisfy

$$\lambda_0 + \lambda_1 + \dots + \lambda_{L-1} = 1 \quad (14)$$

and, as before,

$$[\alpha]_k = \{\mathcal{S}_k \in \mathcal{P}_k : \mathcal{S}_k \ni \alpha\} \quad (15)$$

is the set of strategies of \mathcal{A} that are similar to α at level k .

In words, conditioning on the degree of similarity between strategies, the revising agent considers all possible switches – i.e., to *any* strategy, irrespective of similarity with the agent’s incumbent strategy – with conditional probability λ_0 , all switches to a strategy with at least one degree of similarity with conditional probability λ_1 , and more generally, all strategies with at least k degrees of similarity with conditional probability λ_k . Specifically, in our running example, an agent taking the bus to work would first contemplate whether they want to only consider switching to another bus line (2 degrees of similarity), or to some other means of public transportation (at least 1 degree of similarity), or to consider all possible means of commuting from home to work (i.e., with no regard for similarities).

- (2) Conditioned on the above choice of $k \in \{0, \dots, L-1\}$, the revising α -strategist meets another individual – the comparator – drawn uniformly at random from the sub-population of agents playing a strategy from the chosen similarity class $[\alpha]_k$. Hence, if the population is at state $x \in \mathcal{X}$, the probability that a revising α -strategist encounters a β -strategist is

$$\begin{aligned} P_{\alpha\beta}^{\text{meet}}(x) &= \lambda_0 x_\beta + \lambda_1 x_\beta \frac{\mathbb{1}\{\beta \overset{1}{\sim} \alpha\}}{\sum_{\alpha' \overset{1}{\sim} \alpha} x_{\alpha'}} + \dots + \lambda_{L-1} x_\beta \frac{\mathbb{1}\{\beta \overset{L-1}{\sim} \alpha\}}{\sum_{\alpha' \overset{L-1}{\sim} \alpha} x_{\alpha'}} \\ &= \sum_{k=0}^{L-1} \lambda_k x_{\beta|[\alpha]_k} \end{aligned} \quad (16)$$

where, as per (17), the notation

$$x_{\beta|[\alpha]_k} = \mathbb{1}\{\beta \overset{k}{\sim} \alpha\} \frac{x_\beta}{x_{[\alpha]_k}} \quad (17)$$

stands for the relative frequency of β -strategists within the sub-population of $[\alpha]_k$ -type strategists (i.e., the part of the population playing a strategy with at least k degrees of similarity to α).

- (3) Finally, as in the case of (PPI), the revising agent observes the comparator's payoff: if it is higher than their own, the revising agent switches to the comparator's strategy with probability proportional to the observed payoff excess; otherwise, the revising agent skips the opportunity to switch.

Hence, putting together all of the above yields the *nested pairwise proportional imitation* protocol

$$\varrho_{\alpha\beta}(\pi, x) = P_{\alpha\beta}^{\text{meet}}(x) \cdot [\pi_\beta - \pi_\alpha]_+ = \sum_{k=0}^{L-1} \lambda_k x_{\beta|[\alpha]_k} [\pi_\beta - \pi_\alpha]_+, \quad (\text{NPPI})$$

where, to avoid trivialities, we assume that $\lambda_0 > 0$ (so all strategies are sampled with positive probability). Then, to derive the induced dynamics, a direct substitution into (MD) gives

$$\dot{x}_\alpha = \sum_{k=0}^{L-1} \lambda_k x_\alpha \left[\sum_{\beta \in \mathcal{A}} x_{\beta|[\alpha]_k} [v_\alpha(x) - v_\beta(x)]_+ - \sum_{\beta \in \mathcal{A}} x_{\beta|[\alpha]_k} [v_\beta(x) - v_\alpha(x)]_+ \right] \quad (18)$$

so, after rearranging, we obtain the *nested replicator dynamics*

$$\dot{x}_\alpha = x_\alpha \sum_{k=0}^{L-1} \lambda_k [v_\alpha(x) - u_{[\alpha]_k}(x)] \quad (\text{NRD})$$

where

$$u_{\mathcal{S}}(x) = \sum_{\alpha \in \mathcal{S}} x_\alpha v_\alpha(x) = \frac{1}{x_{\mathcal{S}}} \sum_{\alpha \in \mathcal{S}} x_\alpha v_\alpha(x) \quad (19)$$

denotes the mean payoff of the class $\mathcal{S} \subseteq \mathcal{A}$. The protocol (NPPI) and the dynamics (NRD) will be our main object of study, so we close this section with some remarks intended to highlight their most salient features.

Remark 1. We begin by noting that (PPI) and (RD) are immediately recovered from (NPPI) and (NRD) respectively by taking $\lambda_0 = 1$ and $\lambda_1 = \dots = \lambda_{L-1} = 0$. In this case, revising agents do not take into account similarities between strategies when selecting a comparator, so it stands to reason that the induced dynamics boil down to (RD). §

Remark 2. Moving forward, it is important to note that comparison probabilities under (NPPI) are always higher between similar strategies compared to dissimilar ones, and the closer the similarity, the higher the probability of carrying out a comparison. To make this precise, define the *degree of similarity* between α and β as

$$\deg(\alpha, \beta) := \max\{k = 0, \dots, L-1 : \alpha \stackrel{k}{\sim} \beta\} \quad (20)$$

i.e., α and β are similar at level $\deg(\alpha, \beta)$ but not at level $\deg(\alpha, \beta) + 1$. Then the probability that β is considered by an α -strategist as a candidate for imitation is

$$P_{\alpha\beta}^{\text{cand}} = \mathbb{P}(\beta \in \text{cand}(\alpha)) = \lambda_0 + \lambda_1 + \dots + \lambda_{\deg(\alpha, \beta)} \quad (21)$$

so $P_{\alpha\beta}^{\text{cand}}$ is increasing in $\deg(\alpha, \beta)$, just like the encounter probability $P_{\alpha\beta}^{\text{meet}}$ of (16).

On the other hand, the agents' imitation rates under (NPPI) also depend on the payoff excess $[\pi_\beta - \pi_\alpha]_+$ of β over α ; more precisely, by (17) we have

$$r_{\alpha\beta}(\pi, x) = \frac{\varrho_{\alpha\beta}(\pi, x)}{x_\beta} = \frac{P_{\alpha\beta}^{\text{meet}}(x)}{x_\beta} \cdot [\pi_\beta - \pi_\alpha]_+ = [\pi_\beta - \pi_\alpha]_+ \cdot \sum_{k=0}^{\deg(\alpha, \beta)} \frac{\lambda_k}{x_{[\alpha]_k}} \quad (22)$$

so the total imitation rate is not always higher for more similar strategies (after all, a less familiar strategy could still be a favorable candidate for imitation, even if it is encountered with lower probability). However, for any *given* value of $[\pi_\beta - \pi_\alpha]_+$, the imitation rate of

β -strategists by α -strategists is highest when α and β are *most* similar (e.g., different metro lines), and lowest when α and β are *least* similar (e.g., different transport modes altogether). This increased affinity toward similar strategies is an *intrinsic* feature of the agents' revision mechanism and it is due to the fact that, ceteris paribus, an agent is more likely to compare strategies with another agent if the other agent's strategy is more similar to their own. §

Remark 3. Building on the previous remark, it is also natural to ask what happens if a revising agent meets another agent from the population uniformly at random (that is, $P_{\alpha\beta}^{\text{meet}}(x) = x_\beta$) but is *extrinsically* more likely to imitate strategies that are more similar to their own. In this case, the agents' imitation rates can be written as

$$r_{\alpha\beta}(\pi, x) = \Theta_{\alpha\beta} \cdot \frac{P_{\alpha\beta}^{\text{meet}}(x)}{x_\beta} \cdot [\pi_\beta - \pi_\alpha]_+ = \Theta_{\alpha\beta} [\pi_\beta - \pi_\alpha]_+ \quad (23)$$

where the imitation coefficients $\Theta_{\alpha\beta}$ are of the form

$$\Theta_{\alpha\beta} = \theta_0 + \theta_1 + \dots + \theta_{\deg(\alpha, \beta)} \quad (24)$$

for some $\theta_0, \dots, \theta_{L-1} \geq 0$ (so $\Theta_{\alpha\beta}$ increases with the degree of similarity between α and β).

The expression (23) for $r_{\alpha\beta}$ inflates the probability of imitating a similar strategy by a factor of $\Theta_{\alpha\beta}$, so the revising agent is more likely to imitate the strategy of another agent if the other agent's strategy is more similar to their own. Mutatis mutandis, this observation also applies to the agents' imitation rates (22) under (NPPI), but where the two models differ is that, in (22), the reinforcement of similar strategies arises *intrinsically*, as a consequence of the fact that an agent is more likely to compare strategies with another agent if the other agent's strategy is more similar to their own. By contrast, the reinforcement of similar strategies in (23) is *extrinsic*, and it occurs *despite* the fact that players meet other players uniformly at random.

The difference between these two models – intrinsic vs. extrinsic – can also be seen in the induced dynamics. Indeed, substituting $\varrho_{\alpha\beta} = r_{\alpha\beta}x_\beta$ in (MD) and rearranging yields the dynamics

$$\dot{x}_\alpha = x_\alpha \sum_{k=0}^{L-1} \theta_k x_{[\alpha]_k} [v_\alpha(x) - u_{[\alpha]_k}(x)] \quad (\text{NRD}_{\text{extr}})$$

where the subscript “extr” alludes to the extrinsic nature of the model (23). Compared to (NRD), the main factor setting $(\text{NRD}_{\text{extr}})$ apart is that the state-independent coefficients λ_k in the former are replaced by the state-dependent factors $\theta_k x_{[\alpha]_k}$ in the latter. Since $x_{[\alpha]_k} \leq 1$, this adjustment means that the impact of the intra-class revision terms $v_\alpha(x) - u_{[\alpha]_k}(x)$ is toned down in $(\text{NRD}_{\text{extr}})$ relative to (NRD). In turn, this reflects the fact that comparisons to similar strategies occur less frequently in the extrinsic model (23) rather than the intrinsic model (22), so the relative ranking of strategies within a similarity class is likewise less impactful in $(\text{NRD}_{\text{extr}})$ than it is in (NRD).

The extrinsic dynamics $(\text{NRD}_{\text{extr}})$ are clearly interesting in their own right, but a full analysis would take us too far afield, so we do not undertake it here. §

Remark 4. It is also worth noting that summing (NRD) over all alternatives in a given similarity class $\mathcal{S} \in \mathcal{P}_\ell$ yields the similarity class dynamics

$$\dot{x}_\mathcal{S} = \sum_{\alpha \in \mathcal{S}} \dot{x}_\alpha = x_\mathcal{S} \sum_{k=0}^{\ell-1} \lambda_k [u_\mathcal{S}(x) - u_{[\mathcal{S}]_k}(x)] \quad (25)$$

where, as before, $[\mathcal{S}]_k$ denotes the k -th level ancestor of \mathcal{S} . These dynamics are formally analogous to (NRD) with ℓ levels of nesting instead of L , and with a coarser similarity hierarchy that does not differentiate actions beyond ℓ degrees of similarity. However, the

average payoffs $u_S(x)$ and $u_{[\alpha]_k}(x)$ appearing in (25) depend on the *complete* population state $x \in \mathcal{X}$, not only on the aggregate population shares x_S . Because of this, the evolution of the population cannot be described by (25) alone if $\ell < L$, i.e., (25) does not stand on its own as an autonomous dynamical system. §

Remark 5. As a final point, we note that the per capita growth rate $g_\alpha(x) = \dot{x}_\alpha/x_\alpha$ of $\alpha \in \mathcal{A}$ under (NRD) is

$$g_\alpha(x) = \sum_{k=0}^{L-1} \lambda_k [v_\alpha(x) - u_{[\alpha]_k}(x)] = v_\alpha(x) - \sum_{k=0}^{L-1} \lambda_k u_{[\alpha]_k}(x) \quad (26)$$

i.e., $g_\alpha(x)$ is a convex combination of payoff differences between $v_\alpha(x)$ and the sub-population means $u_{[\alpha]_k}(x)$ evaluated at all levels of resemblance to α , from the maximum possible degree (when $k = L - 1$) to no similarities whatsoever (when $k = 0$). An important consequence of this is that, for any pair of strategies $\alpha, \beta \in \mathcal{A}$, we have

$$g_\alpha(x) - g_\beta(x) = v_\alpha(x) - v_\beta(x) - \sum_{k=0}^{L-1} \lambda_k [u_{[\alpha]_k}(x) - u_{[\beta]_k}(x)]. \quad (27)$$

Clearly, if α and β are *siblings* – that is, children of the same parent, so $[\alpha]_k = [\beta]_k$ for all $k = 0, \dots, L - 1$ – Eq. (27) gives $g_\alpha(x) - g_\beta(x) = v_\alpha(x) - v_\beta(x)$, so the dynamics (NRD) exhibit payoff-monotone growth rates within the *finest* non-trivial similarity class.¹² On the other hand, if α and β are not siblings, the convex combination of class payoff differences in (27) could be arbitrarily large relative to $v_\alpha(x) - v_\beta(x)$, and with a different sign. Thus, in stark contrast to (RD), the nested dynamics (NRD) are *not payoff-monotonic*, a fact which poses significant obstacles to their analysis. We will revisit this point in the next section. §

4. RATIONALITY ANALYSIS

4.1. Statement of results and discussion. We are now in a position to state and prove our main results concerning the long-run rationality properties of (NRD). These are as follows:¹³

Theorem 1. *Let $\mathcal{G} \equiv \mathcal{G}(\mathcal{A}, v)$ be a population game, let \mathcal{P}_\bullet be a similarity structure on \mathcal{A} , and let $x(t)$ be an interior solution of the nested replicator dynamics (NRD) for \mathcal{G} . Then:*

1. *If $\alpha \in \mathcal{A}$ is strictly dominated, then $\lim_{t \rightarrow \infty} x_\alpha(t) = 0$.*
2. *A state $x^* \in \mathcal{X}$ is stationary if and only if it is a restricted equilibrium of \mathcal{G} .*
3. *If $x(t)$ converges to x^* as $t \rightarrow \infty$, then x^* is a Nash equilibrium of \mathcal{G} .*
4. *If $x^* \in \mathcal{X}$ is Lyapunov stable, then it is a Nash equilibrium of \mathcal{G} .*
5. *If x^* is an ESS of \mathcal{G} , it is asymptotically stable. In particular, if x^* is a GESS of \mathcal{G} , it attracts all initial conditions with $\text{supp}(x(0)) \supseteq \text{supp}(x^*)$.*
6. *If \mathcal{G} is a potential game with potential function F , then $F(x(t))$ is nondecreasing and $x(t)$ converges to a connected set of restricted equilibria of \mathcal{G} .*

¹²Following Sandholm (2010a, 2015) and Weibull (1995), payoff monotonicity means here that $g_\alpha(x) \geq g_\beta(x)$ if and only if $v_\alpha(x) \geq v_\beta(x)$. In view of (22), Weibull (1995) would call (NPPI) an “imitative” protocol, but this clashes with Sandholm (2010a, Chap. 5.4), where a payoff monotonicity condition is appended to the definition of “imitative”; for a detailed discussion of this issue, see Mertikopoulos & Viossat (2022).

¹³We recall here very briefly that $x^* \in \mathcal{X}$ is said to be (i) *Lyapunov stable* (or simply *stable*) if every solution trajectory $x(t)$ that starts close enough to x^* remains close enough to x^* for all $t \geq 0$; (ii) *attracting* if $\lim_{t \rightarrow \infty} x(t) = x^*$ for every solution trajectory $x(t)$ that starts close enough to x^* ; and (iii) *asymptotically stable* if it is both stable and attracting. For an introduction to the theory of dynamical systems, we refer the reader to Shub (1987) and Hirsch et al. (2004).

7. If \mathcal{G} is strictly monotone, its (necessarily unique) Nash equilibrium attracts almost all initial conditions – and, in particular, all initial conditions with full support.

Before discussing the proof of [Theorem 1](#), it is worth noting that, as we explained in the previous section, the dynamics ([NRD](#)) are *not* payoff-monotonic. On that account, it is fairly surprising to see that they satisfy a range of properties that have long been associated with payoff monotonicity (such as the elimination of strictly dominated strategies and the stability properties of Nash equilibria).

To illustrate the intricacies that this entails, consider the simple case where α is strictly dominated by β with $v_\beta(x) - v_\alpha(x) = c > 0$ for all $x \in \mathcal{X}$. Then, by [\(27\)](#), we readily get

$$\frac{d}{dt} \log \frac{x_\alpha}{x_\beta} = \frac{\dot{x}_\alpha}{x_\alpha} - \frac{\dot{x}_\beta}{x_\beta} = -c - \sum_{k=0}^{L-1} \lambda_k [u_{[\alpha]_k}(x) - u_{[\beta]_k}(x)]. \quad (28)$$

Now, if α and β are similarity siblings (that is, $\alpha \stackrel{L-1}{\sim} \beta$), we will have $[\alpha]_k = [\beta]_k$ for all $k = 0, \dots, L-1$, so [\(28\)](#) gives

$$x_\alpha(t) \leq \exp(b - ct) \quad (29)$$

for some constant $b \in \mathbb{R}$ depending on the initialization of ([NRD](#)). Thus, since [\(29\)](#) can be tight in the limit $t \rightarrow \infty$ if $x_\beta(t) \rightarrow 1$, we conclude that α becomes extinct in the long run and the worst-case rate of extinction is $\Theta(\exp(-ct))$.

On the other hand, if α and β only share, say, $L-2$ degrees of similarity (e.g., they are different means of transportation but both public, like a bus line and a metro line in our running example), [Eq. \(28\)](#) instead yields

$$\frac{d}{dt} \log \frac{x_\alpha}{x_\beta} = -c - \lambda_{L-1} [u_{[\alpha]_{L-1}}(x) - u_{[\beta]_{L-1}}(x)]. \quad (30)$$

In this case, there is no a priori reason to assume that, as a group, the parent class $[\beta]_{L-1}$ of β is doing better in terms of payoffs than the parent class $[\alpha]_{L-1}$ of α (for instance, it may well happen that a certain metro line is dominated by a specific bus line, while all other metro lines dominate all bus lines). Because of this, the class payoff difference $u_{[\alpha]_{L-1}}(x) - u_{[\beta]_{L-1}}(x)$ in [\(30\)](#) could cancel out the domination margin c between β and α , in which case the per capita growth rate of α could even *exceed* the per capita growth rate of β , although α is dominated by β .

In view of the above, the fact that dominated strategies become extinct under ([NRD](#)) irrespective of similarities may appear somewhat puzzling – and, perhaps, unexpected. Nonetheless, the degree of similarity between strategies plays an important role in the dynamics; as we show below, it controls *the rate of extinction* of dominated strategies.

Proposition 1. *Let $x(t)$ be an interior solution of ([NRD](#)), and suppose that $\alpha \in \mathcal{A}$ is dominated by $\beta \in \mathcal{A}$ with a margin of c , that is,*

$$c = \min_{x \in \mathcal{X}} [v_\beta(x) - v_\alpha(x)] > 0. \quad (31)$$

Then α becomes extinct along $x(t)$ at a rate of

$$x_\alpha(t) \leq \exp\left(b - \sum_{k=0}^{\ell} \lambda_k \cdot ct\right) \quad (32)$$

where $\ell = \deg(\alpha, \beta) \equiv \max\{k = 0, \dots, L-1 : \alpha \stackrel{k}{\sim} \beta\}$ is the degree of similarity between α and β , and $b \in \mathbb{R}$ is a constant depending on $x(0)$.

To streamline our presentation, we defer the proof of [Proposition 1](#) to [Appendix A](#). Instead, for our purposes, it is more important to note that the higher the degree of similarity between a dominated strategy and the strategy that dominates it, the faster the rate of

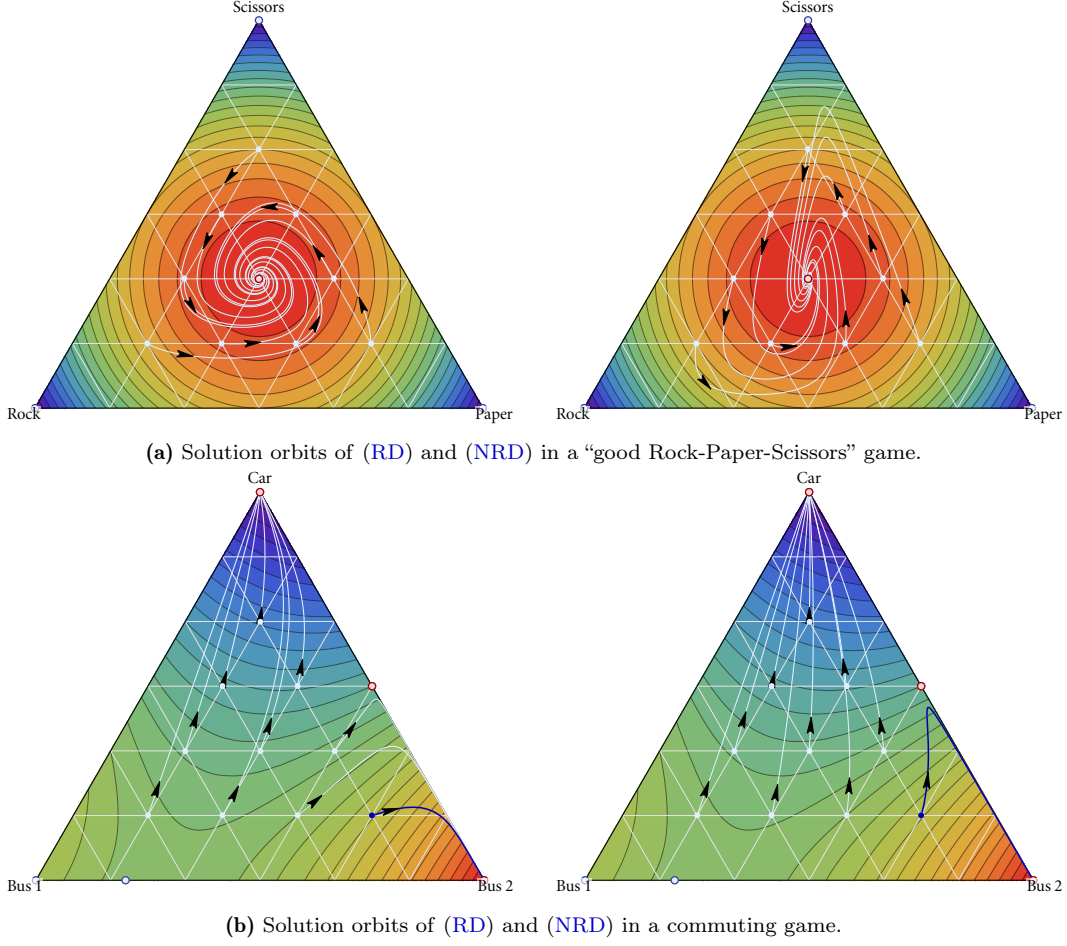


Figure 3: Solution orbits of the replicator dynamics (left) and the nested replicator dynamics (right) in a game of “good Rock-Paper-Scissors” (top) and the commuting game described in the text (bottom). Nash equilibria are depicted in red, stationary points in blue; the contours represent the population mean payoff (higher values shifted to red). In the RPS game, “P” has been arbitrarily grouped with “S” to illustrate the distortion incurred by the implicit similarity bias of (NRD). In all cases, (NRD) has been run with intra-level sampling probabilities $\lambda_0 = 1/4$ and $\lambda_1 = 3/4$. The evolution of the population shares of the highlighted orbit in the commuting game is shown in Fig. 4.

extinction of said strategy. This is due to the revision mechanism that is driving the dynamics (NRD): because an agent is less likely to initiate a comparison with less similar strategies, the probability of dropping a dominated strategy is likewise reduced if the strategy that dominates it has a small degree of similarity. This decrease is captured by the coefficient $\Lambda_\ell \equiv \lambda_0 + \dots + \lambda_\ell$ that modulates the domination margin c in (32), and which is strictly less than 1 if α and β are similar to the highest possible degree.

Remark. A similar result can be obtained for the *rate of convergence* of (NRD) to strict equilibria. In particular, if x^* is a strict Nash equilibrium of \mathcal{G} and $c = \min_{\beta \notin \text{supp}(x^*)} [u(x^*) - v_\beta(x^*)]$ denotes the minimum payoff difference at equilibrium, it can be shown that, for all

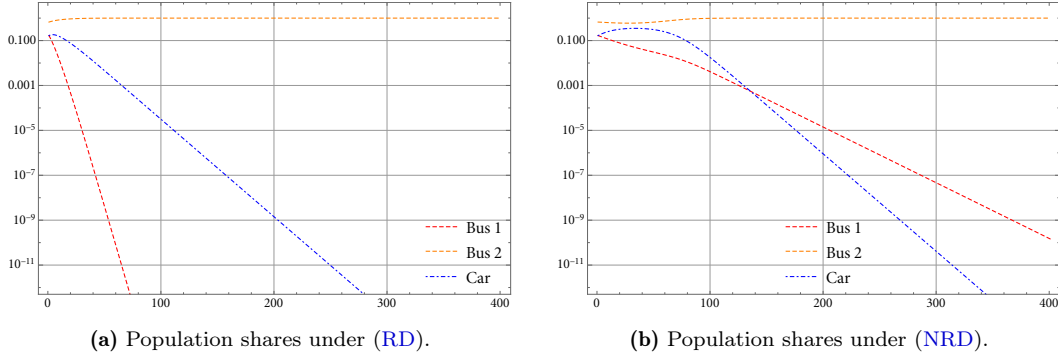


Figure 4: The rate of extinction of dominated strategies over time under the replicator dynamics (left) and the nested replicator dynamics (right). Population shares are computed for the highlighted orbits of Fig. 3; axes are log-linear, indicating an exponential rate of convergence to equilibrium, and an exponential rate of extinction of dominated strategies (with the slope of each line capturing the extinction exponent). In tune with Proposition 1, we observe that the first bus line becomes extinct at a significantly faster rate when similarities are not taken into account.

$\varepsilon > 0$, there exists a neighborhood \mathcal{U} of x^* in \mathcal{X} such that

$$\|x(t) - x^*\|_1 \leq \exp\left(b - (1 - \varepsilon) \sum_{k=0}^{\ell} \lambda_k \cdot ct\right) \quad \text{whenever } x(0) \in \mathcal{U} \quad (33)$$

where ℓ denotes the degree of similarity between x^* and the strategy $\beta \notin \text{supp}(x^*)$ that realizes the value of c . Up to the factor $1 - \varepsilon$ (which is due to technical reasons, and which can be taken arbitrarily close to 1), we observe that the degree of similarity has the same effect here as it does on the rate of extinction of dominated strategies. \S

As a concrete illustration of the above, consider the following commuting game (loosely based on the blue bus / red bus paradox). A population of commuters can choose between taking either of two bus lines (labeled as “1” and “2”) or their car (strategy “3”) to go to work; going by car is faster when congestion is light, but taking the bus becomes the better choice as car traffic increases. We represent this via the affine payoff model $v(x) = Mx + b$ with

$$b = -\begin{pmatrix} 5 \\ 5 \\ 2 \end{pmatrix} \quad \text{and} \quad M = -\begin{pmatrix} 2 & 4 & 6 \\ 3 & 0 & 6 \\ 1 & 4 & 8 \end{pmatrix}. \quad (34)$$

A direct calculation (which we omit) shows that this game admits three Nash equilibria, two strict, $(0, 1, 0)$ and $(0, 0, 1)$, and one mixed, $(0, 1/2, 1/2)$. In addition, strategy 1 (the first bus line) is dominated by strategy 3 (commuting by car) with a margin of $c = 1$, and there are no other dominated strategies.

To demonstrate the impact of similarities, we depict in Fig. 3 the solution orbits of the replicator dynamics (RD) and the nested replicator dynamics (NRD) for this game (with the two bus lines regarded as “similar” by the commuters with an intra-class sampling probability of $\lambda_1 = 3/4$). In tune with Theorem 1, the stability characteristics of the game’s Nash equilibria do not change: the two strict equilibria (which are evolutionarily stable) remain asymptotically stable under both (RD) and (NRD), while the game’s mixed equilibrium is unstable. From a quantitative viewpoint however, the basins of attraction of the game’s stable

equilibria are quite different in the nested and the non-nested case: specifically, commuting by car attracts a significantly larger set of initial conditions when the two bus lines are regarded as similar by the population. This is also reflected in the rate of extinction of dominated strategies: as we show in Fig. 4, the rate of extinction of the first bus line is significantly slower when bus users are more likely to compare strategies with other bus users, precisely because neither bus line dominates the other. This is captured by the coefficient $\sum_{k=0}^{\ell} \lambda_k$ of (32): in the case of (RD) it is equal to 1 (because similarities are not taken into account), whereas in the case of (NRD) it is equal to λ_0 and thus, potentially much smaller.

4.2. Proof outline. We close this section with an outline of the proof of Theorem 1 and Proposition 1 (full proofs are provided in Appendix A). In contrast to standard arguments relying on payoff monotonicity, the enabling machinery for the rationality properties of (NRD) stated in Theorem 1 requires the construction of a specific “energy function” generalizing the well-known Kullback–Leibler (KL) divergence

$$D(p, x) = \sum_{\alpha \in \mathcal{A}} p_{\alpha} \log \frac{p_{\alpha}}{x_{\alpha}}. \quad (35)$$

With a fair degree of hindsight, we will instead consider the *nested Kullback–Leibler divergence*

$$D_L(p, x) = \sum_{\ell=1}^L \delta_{\ell} \sum_{S \in \mathcal{P}_{\ell}} p_S \log \frac{p_S}{x_S} \quad (\text{NKL})$$

where the weights δ_{ℓ} , $\ell = 1, \dots, L$, are given by the somewhat mysterious expression

$$\delta_{\ell} = \begin{cases} \lambda_{\ell}/(\Lambda_{\ell-1}\Lambda_{\ell}) & \text{if } \ell = 1, \dots, L-1, \\ 1 & \text{if } \ell = L, \end{cases} \quad (36)$$

with $\Lambda_{\ell} = \lambda_0 + \dots + \lambda_{\ell}$ denoting the total intra-class revision rate at level ℓ and below (recall here our normalization assumption $\Lambda_{L-1} = \lambda_0 + \dots + \lambda_{L-1} = 1$).

Intuitively, D_L can be viewed as an affine combination of the ordinary KL divergence between p and x at different degrees of similarity: in particular, the ℓ -th level terms in (NKL) add up to the standard KL divergence between the class state variables $(p_{S_{\ell}})_{S_{\ell} \in \mathcal{P}_{\ell}}$ and $(x_{S_{\ell}})_{S_{\ell} \in \mathcal{P}_{\ell}}$ in $\Delta(\mathcal{P}_{\ell})$. As for the – admittedly opaque – choice of weights δ_k in the definition of D_L , the specific assignment (36) has been judiciously chosen to provide the following fundamental link between the nested divergence (NKL) and the nested dynamics (NRD):

Proposition 2. *Fix a state $p \in \mathcal{X}$ and let $x(t)$ be an interior solution of (NRD). Then:*

$$\frac{d}{dt} D_L(p, x(t)) = \langle v(x(t)), x(t) - p \rangle = \sum_{\alpha \in \mathcal{A}} v_{\alpha}(x(t)) (x_{\alpha}(t) - p_{\alpha}) \quad (37)$$

In words, Proposition 2 states that, under (NRD), the rate of change of the nested KL divergence relative to a base point $p \in \mathcal{X}$ is the difference between the population average payoff $u(x) = \langle v(x), x \rangle$ and the average payoff $\langle v(x), p \rangle$ of a group of mutants with actions distributed according to p (for a range of similar results in the context of game dynamics, see Mertikopoulos & Sandholm, 2016, Mertikopoulos & Zhou, 2019, and references therein). The key property highlighted by this interpretation is that two similarity-*dependent* objects – the nested dynamics (NRD) and the nested KL divergence (NKL) – give rise to a similarity-*agnostic* quantity that only depends on the game and the population states involved, and is otherwise oblivious of the similarity hierarchy \mathcal{P}_{\bullet} .

The main ingredient of the proof of Proposition 2 is that, under the specific weight assignment (36), the time derivative of the nested divergence D_L becomes a telescoping sum,

with only the finest ($\ell = L$) and coarsest ($\ell = 0$) levels surviving in the right-hand side of (37). This requires some delicate algebraic manipulations, which we detail below.

Proof of Proposition 2. By the definition of the nested KL divergence, we have

$$\frac{d}{dt} D_L(p, x(t)) = - \sum_{\ell=1}^L \delta_\ell \sum_{\mathcal{S} \in \mathcal{P}_\ell} p_{\mathcal{S}} \frac{\dot{x}_{\mathcal{S}}}{x_{\mathcal{S}}}. \quad (38)$$

Thus, by substituting the class dynamics (25) into (38) and rearranging, we get

$$\begin{aligned} \frac{d}{dt} D_L(p, x(t)) &= - \sum_{\ell=1}^L \delta_\ell \sum_{\mathcal{S} \in \mathcal{P}_\ell} p_{\mathcal{S}} \sum_{k=0}^{\ell-1} \lambda_k [u_{\mathcal{S}}(x) - u_{[\mathcal{S}]_k}(x)] \\ &= \sum_{\ell=1}^L \delta_\ell \sum_{k=0}^{\ell-1} \lambda_k \sum_{\mathcal{S} \in \mathcal{P}_\ell} p_{\mathcal{S}} [u_{\mathcal{S}}(x) - u_{[\mathcal{S}]_k}(x)]. \end{aligned} \quad (39)$$

To proceed, note that, for all $k \leq \ell$, we have

$$\begin{aligned} \sum_{\mathcal{S} \in \mathcal{P}_\ell} p_{\mathcal{S}} u_{[\mathcal{S}]_k}(x) &= \sum_{\mathcal{S}_k \in \mathcal{P}_k} \sum_{\mathcal{S} \prec_\ell \mathcal{S}_k} p_{\mathcal{S}} u_{[\mathcal{S}]_k}(x) \\ &= \sum_{\mathcal{S}_k \in \mathcal{P}_k} u_{\mathcal{S}_k}(x) \sum_{\mathcal{S} \prec_\ell \mathcal{S}_k} p_{\mathcal{S}} = \sum_{\mathcal{S}_k \in \mathcal{P}_k} u_{\mathcal{S}_k}(x) p_{\mathcal{S}_k} = \langle v(x), p \rangle_k \end{aligned} \quad (40)$$

where the notation $\langle v(x), p \rangle_k$ stands for

$$\langle v(x), p \rangle_k := \sum_{\mathcal{S}_k \in \mathcal{P}_k} p_{\mathcal{S}_k} u_{\mathcal{S}_k}(x) \quad \text{for all } p, x \in \mathcal{X} \text{ and all } k = 0, \dots, L, \quad (41)$$

and we used the facts that $\bigcup\{\mathcal{S} : \mathcal{S} \prec_\ell \mathcal{S}_k\} = \mathcal{S}_k$ and $\bigcap\{\mathcal{S} : \mathcal{S} \prec_\ell \mathcal{S}_k\} = \emptyset$ whenever $k \leq \ell$. In this way, (39) can be unpacked one attribute at a time to yield

$$\begin{aligned} (39) &= \sum_{\ell=1}^L \delta_\ell \sum_{k=0}^{\ell-1} \lambda_k [\langle v(x), p \rangle_\ell - \langle v(x), p \rangle_k] \\ &= \delta_1 \lambda_0 \langle v(x), p \rangle_0 - \delta_1 \lambda_0 \langle v(x), p \rangle_1 \\ &\quad + \delta_2 [\lambda_0 \langle v(x), p \rangle_0 + \lambda_1 \langle v(x), p \rangle_1] - \delta_2 (\lambda_0 + \lambda_1) \langle v(x), p \rangle_2 \\ &\quad \vdots \\ &\quad + \delta_L [\lambda_0 \langle v(x), p \rangle_0 + \dots + \lambda_{L-1} \langle v(x), p \rangle_{L-1}] - \delta_L (\lambda_0 + \dots + \lambda_{L-1}) \langle v(x), p \rangle_L \end{aligned} \quad (42)$$

and hence, regrouping all same-level payoff terms, we obtain:

$$\begin{aligned} \frac{d}{dt} D_L(p, x(t)) &= \lambda_0 (\delta_1 + \dots + \delta_L) \cdot \langle v(x), p \rangle_0 \\ &\quad \vdots \\ &\quad + [\lambda_k (\delta_{k+1} + \dots + \delta_L) - \delta_k (\lambda_0 + \dots + \lambda_{k-1})] \cdot \langle v(x), p \rangle_k \\ &\quad \vdots \\ &\quad - \delta_L (\lambda_0 + \dots + \lambda_{L-1}) \cdot \langle v(x), p \rangle_L \\ &= \sum_{k=0}^L \left[\lambda_k \sum_{\ell=k+1}^L \delta_\ell - \delta_k \sum_{\ell=0}^{k-1} \lambda_\ell \right] \langle v(x), p \rangle_k \end{aligned} \quad (43)$$

where, in the last line, we used the empty sum convention $\sum_{j \in \emptyset} \alpha_j = 0$.

To proceed, we claim that

$$\delta_{k+1} + \dots + \delta_L = \frac{1}{\lambda_0 + \dots + \lambda_k} = \frac{1}{\Lambda_k} \quad \text{for all } k = 0, \dots, L-1. \quad (44)$$

Indeed, for $k = L-1$, we recover the definition (36) of δ_L . Otherwise, assume inductively that (44) holds for some $k < L-1$; then, by (36), we get

$$\delta_k + \delta_{k+1} + \dots + \delta_L = \frac{\lambda_k}{\Lambda_k \Lambda_{k-1}} + \frac{1}{\Lambda_k} = \frac{\lambda_k + \Lambda_{k-1}}{\Lambda_k \Lambda_{k-1}} = \frac{\Lambda_k}{\Lambda_k \Lambda_{k-1}} = \frac{1}{\Lambda_{k-1}}, \quad (45)$$

which completes the inductive step and proves our claim. Hence, combining (44) with (36), we get

$$\lambda_k(\delta_{k+1} + \dots + \delta_L) = \frac{\lambda_k}{\Lambda_k} = \delta_k \Lambda_{k-1} = \delta_k(\lambda_0 + \dots + \lambda_{k-1}) \quad (46)$$

for all $k = 1, \dots, L-1$. Thus, substituting (44) and (46) back into (43), we finally obtain

$$\frac{d}{dt} D_L(p, x(t)) = \langle v(x), p \rangle_0 - \langle v(x), p \rangle_L. \quad (47)$$

However, since the only 0-th level similarity class is \mathcal{A} itself (and $p_{\mathcal{A}} = 1$), we readily get

$$\langle v(x), p \rangle_0 = p_{\mathcal{A}} u_{\mathcal{A}}(x) = \sum_{\alpha \in \mathcal{A}} x_{\alpha} v_{\alpha}(x) = \langle v(x), x \rangle \quad (48)$$

and our assertion follows. \blacksquare

Proposition 2 plays a crucial role in establishing the long-run rationality properties of (NRD) recorded in **Theorem 1**. As an illustration, we sketch below the proof of the theorem's first item concerning the elimination of strictly dominated strategies. For this, suppose that action $\alpha \in \mathcal{A}$ is strictly dominated by action $\beta \in \mathcal{A}$. Then, if we let $e_{\alpha} \in \mathcal{X}$ denote the “pure” population state corresponding to α (and likewise for β), the difference

$$D_L(e_{\alpha}, x(t)) - D_L(e_{\beta}, x(t)) = \sum_{\ell=1}^L \delta_{\ell} \log \frac{x_{[\beta]_{\ell}}(t)}{x_{[\alpha]_{\ell}}(t)}, \quad (49)$$

is a weighted sum of log ratios of the masses of agents using actions in classes containing α and β . It then follows from **Proposition 2** that

$$\begin{aligned} \frac{d}{dt} (D_L(e_{\alpha}, x(t)) - D_L(e_{\beta}, x(t))) &= \langle v(x(t)), x(t) - e_{\alpha} \rangle - \langle v(x(t)), x(t) - e_{\beta} \rangle \\ &= v_{\beta}(x(t)) - v_{\alpha}(x(t)). \end{aligned} \quad (50)$$

Since β strictly dominates α , the payoff difference (50) is positive and bounded away from zero. Because of this, the difference (49) must grow without bound, which in turn implies that the population share x_{α} of the dominated strategy α vanishes, and a more detailed algebraic manipulation allows us to get the rate estimate described in **Proposition 1**.

For a detailed proof of the above claims and the remaining parts of **Theorem 1** and **Proposition 1**, we refer the reader to **Appendix A**.

5. EVOLUTION THROUGH REINFORCEMENT AND NESTED LOGIT CHOICE

In this section, we discuss a continuous-time model for learning in games, which reveals a deep relation between the nested replicator dynamics and the nested logit choice rule of **Ben-Akiva (1973)** and **McFadden (1978)**. To establish a baseline, we begin by describing the non-nested case in **Section 5.1** below; subsequently, we present the details of this relation in full in **Sections 5.2** and **5.3** right after.

5.1. Logit choice, exponential weights, and the replicator dynamics. Our starting point is the following stimulus-response model of evolution in the spirit of [Erev & Roth \(1998\)](#):

- At each instance $t \geq 0$, the agents observe the cumulative payoff $y_\alpha(t)$ of each pure strategy $\alpha \in \mathcal{A}$ up to time t , viz.

$$y_\alpha(t) = \int_0^t v_\alpha(x(\tau)) d\tau \quad (51)$$

- Subsequently, players select actions with probabilities that are exponentially proportional to the aggregate propensity scores y_α , reinforcing in this way those actions that seem to be performing better. As a result, in our nonatomic setting, the relative share of α -strategists at any given time are determined by the exponential choice rule

$$x_\alpha(t) = \frac{\exp(y_\alpha(t)/\mu)}{\sum_{\beta \in \mathcal{A}} \exp(y_\beta(t)/\mu)} \quad (52)$$

where $\mu > 0$ is a “temperature parameter” that indicates the players’ sensitivity to changes in the observed payoffs. When $\mu \rightarrow \infty$, players are relatively indifferent and tend to employ all strategies in a near-uniform mix; by contrast, when $\mu \rightarrow 0$, players are extremely sensitive to payoff changes and essentially “best-respond” to the vector of aggregate scores $y(t)$.

The main difference between this model and the revisionist viewpoint of the previous sections is that, instead of specifying the inflows and outflows to and from each strategy, [Eq. \(52\)](#) directly prescribes the distribution of the population as a function of the cumulative propensity scores [\(51\)](#). Nevertheless, despite these very different starting points, these two processes yield the same evolutionary dynamics: indeed, by expressing [Eqs. \(51\)](#) and [\(52\)](#) in differential form, we obtain the *exponential weights* (EW) dynamics

$$\dot{y} = v(x) \quad x = \Lambda(y) \quad (\text{EW})$$

where

$$\Lambda_\alpha(y) = \frac{\exp(y_\alpha/\mu)}{\sum_{\beta \in \mathcal{A}} \exp(y_\beta/\mu)} \quad (\text{LC})$$

denotes the logit choice map underlying [\(52\)](#). Then, by a straightforward differentiation, we readily get the system

$$\dot{x}_\alpha = \mu^{-1} x_\alpha [v_\alpha(x) - u(x)] \quad (53)$$

which, up to the multiplicative factor μ^{-1} , is simply the replicator equation [\(RD\)](#) induced by the revision protocol [\(PPI\)](#).

Remark. The exponential weights scheme [\(EW\)](#) is one of the staples of the online learning literature and its origins can be traced to the work of [Vovk \(1990\)](#), [Littlestone & Warmuth \(1994\)](#) and [Auer et al. \(1995, 2002\)](#) on multi-armed bandits. The specific derivation of the replicator dynamics that we presented here goes back at least as far as [Rustichini \(1999\)](#) but versions of this model have also been studied in different contexts by [Börgers & Sarin \(1997\)](#), [Posch \(1997\)](#), [Hopkins \(2002\)](#), [Beggs \(2005\)](#), [Leslie & Collins \(2005\)](#), [Hofbauer et al. \(2009\)](#), [Mertikopoulos & Moustakas \(2010\)](#), [Coucheney et al. \(2015\)](#), and many others; for a partial survey, see [Mertikopoulos et al. \(2024\)](#) and references therein.

A key byproduct of this equivalence between [\(RD\)](#) and [\(EW\)](#) is that the latter immediately inherits all the rationality properties of the former, including – but not limited to – [Theorem 1](#). This suite of properties has contributed significantly to the popularity of [\(EW\)](#) in the literature on learning in games, so a natural question that arises is whether a similar link can be established between the nested dynamics [\(NRD\)](#) and a suitable model for learning in the presence of similarities. We address this question in the following sections.

5.2. Nested logit choice. Going back to the derivation of the dynamics (EW), a key observation is that the action selection rule (LC) puts all alternatives on an equal footing and is oblivious to any exogenous similarities they might possess. Albeit reasonable in certain contexts, this feature of (LC) can become problematic when the inclusion of irrelevant alternatives ends up conflating the action choices of the population and lead to undesirable behaviors.

For an illustration, consider the following adaptation of the classical “red bus / blue bus” paradox of McFadden (1974b). To wit, suppose that a population of commuters can take a car or bus to work and, on average, commuting by car takes half as long as commuting by bus. Accordingly, if the induced propensities (51) for employing each alternative are, say, $y_{\text{car}} = -1/2$ and $y_{\text{bus}} = -1$, the ratio of commuters taking a car over a bus under (LC) would be $x_{\text{car}}/x_{\text{bus}} = \exp(-1/2)/\exp(-1) \approx 1.6$ in favor of taking a car, in tune with the fact that commuting by bus takes much longer. The paradox now comes up if the company operating the bus network paints half of the buses blue and the other half red. This change has absolutely no effect on the travel time of a bus, but since the commuters’ new set of alternatives is $\mathcal{A} = \{\text{car}, \text{red bus}, \text{blue bus}\}$, the ratio of commuters taking a car under (LC) would drop to $x_{\text{car}}/[x_{\text{red bus}} + x_{\text{blue bus}}] = \exp(-1/2)/[\exp(-1) + \exp(-1)] \approx 0.8$, i.e., half of what it was before the addition of an irrelevant feature (the color of the bus), and no longer consistent with the associated travel times.

To avoid this artificial inflation of similar alternatives, Ben-Akiva (1973) and McFadden (1974a, 1978) introduced the so-called nested logit choice (NLC) rule, which does not treat all alternatives together at the same time; instead, it singles out an alternative by progressively filtering out similarities within the set of all possible alternatives, one attribute at a time. In more detail, given a similarity structure \mathcal{P}_\bullet and a profile of propensity scores $y = (y_1, \dots, y_m) \in \mathbb{R}^{\mathcal{A}}$ on \mathcal{A} , the nested logit choice model unfolds as follows:

- (1) First, the propensity score of a similarity class $\mathcal{S}_k \in \mathcal{P}_k$, $k = 0, \dots, L-1$, is defined inductively as

$$y_{\mathcal{S}_{\ell-1}} = \mu_\ell \log \sum_{\mathcal{S}_\ell \triangleleft \mathcal{S}_{\ell-1}} \exp\left(\frac{y_{\mathcal{S}_\ell}}{\mu_\ell}\right) \quad (54)$$

where $\mu_\ell > 0$, $\ell = 1, \dots, L$, measures the players’ *uncertainty level* relative to the ℓ -th tier \mathcal{P}_ℓ of \mathcal{P}_\bullet . Thus, starting with the individual alternatives of \mathcal{A} – that is, the leaves $\{\alpha\} \in \mathcal{P}_L$ of \mathcal{P}_\bullet – propensity scores are backpropagated along \mathcal{P}_\bullet , one attribute at a time, until reaching the root $\mathcal{P}_0 = \{\mathcal{A}\}$ of \mathcal{P}_\bullet . The reason for defining the class scores $y_{\mathcal{S}_k}$ in this precise way is discussed in detail below; for now, we only note that (54) defines the score of a class as the weighted softmax of the scores of its children, meaning in particular that a class with higher-scoring options will score higher overall.

- (2) An alternative $\alpha \in \mathcal{A}$ is drawn by specifying a chain of progressively finer similarity classes $\mathcal{A} \equiv \mathcal{S}_0 \triangleright \mathcal{S}_1 \triangleright \dots \triangleright \mathcal{S}_L \equiv \{\alpha\}$, each selected via the family of conditional probabilities

$$P_{\mathcal{S}_\ell | \mathcal{S}_{\ell-1}}(y) = \frac{\exp(y_{\mathcal{S}_\ell} / \mu_\ell)}{\exp(y_{\mathcal{S}_{\ell-1}} / \mu_\ell)} = \frac{\exp(y_{\mathcal{S}_\ell} / \mu_\ell)}{\sum_{\mathcal{S}'_\ell \sim \mathcal{S}_{\ell-1}} \exp(y_{\mathcal{S}'_\ell} / \mu_\ell)} \quad (55)$$

for all $\ell = 1, \dots, L$. In particular, (55) first prescribes choice probabilities $x_{\mathcal{S}_1}$ for all classes $\mathcal{S}_1 \in \mathcal{P}_1$ (i.e., the coarsest ones); subsequently, once a class $\mathcal{S}_1 \in \mathcal{P}_1$ has been selected, (55) prescribes the conditional choice probabilities $x_{\mathcal{S}_2 | \mathcal{S}_1}$ for all children \mathcal{S}_2 of \mathcal{S}_1 and draws a class from \mathcal{P}_2 based on $x_{\mathcal{S}_2 | \mathcal{S}_1}$. The process then continues downward along \mathcal{P}_\bullet until reaching the finest partition \mathcal{P}_L and selecting an alternative $\{\alpha\} \equiv \mathcal{S}_L \in \mathcal{P}_L$.

In this way, unrolling the conditional probabilities (55) over the lineage of any given action $\alpha \in \mathcal{A}$, we obtain the *nested logit choice* model

$$P_\alpha(y) = \prod_{\ell=1}^L \frac{\exp(y_{S_\ell}/\mu_\ell)}{\exp(y_{S_{\ell-1}}/\mu_\ell)} \quad \text{for all } \alpha \in \mathcal{A} \quad (\text{NLC})$$

with each y_{S_ℓ} , $\ell = 1, \dots, L$, given by (54). This action selection rule – originally due to Ben-Akiva (1973) and McFadden (1978) – will be the mainstay of our analysis, so some remarks are in order.

The first thing of note is the precise expression (54) of the class propensities y_S . The reason for this definition is the conditional probability rule (55), which posits that each attribute selection step of (NLC) is an instance of the non-nested choice rule (LC): in particular, treating the parent class $S_{\ell-1}$ as a set of alternatives on its own right, the expression (54) is essentially imposed by the requirement $\sum_{S_\ell \triangleleft S_{\ell-1}} P_{S_\ell|S_{\ell-1}}(y) = 1$. As for the uncertainty parameters $\mu_1, \dots, \mu_L > 0$, these are implicitly assumed to model the level of the agents' uncertainty when selecting a particular attribute: for large values of μ_ℓ , agents tend to be uniformly distributed among the constituent classes of \mathcal{P}_ℓ ; by contrast, when $\mu_\ell \rightarrow 0$, nearly the entire population is employing the similarity class $S_\ell \in \mathcal{P}_\ell$ with the highest aggregate propensity y_{S_ℓ} . Since coarser attributes inherently come with higher levels of uncertainty, we will assume throughout that $\mu_1 \geq \dots \geq \mu_L > 0$; this point will play an important role in the sequel.

Now, other than selecting a primitive alternative $\alpha \in \mathcal{A}$, unrolling the chain of probabilities (55) over the lineage $\mathcal{A} \equiv S_0 \triangleright S_1 \triangleright \dots \triangleright S_\ell \equiv S$ of a target class $S \in \mathcal{P}_\ell$ yields

$$P_S(y) = \frac{\exp(y_S/\mu_\ell)}{\exp(y_{S_{\ell-1}}/\mu_\ell)} \times \dots \times \frac{\exp(y_{S_1}/\mu_1)}{\exp(y_{\mathcal{A}}/\mu_1)} = \prod_{k=1}^\ell \frac{\exp(y_{S_k}/\mu_k)}{\exp(y_{S_{k-1}}/\mu_k)} \quad (56)$$

for the relative share of individuals playing an element of S under the propensity score profile $y \in \mathbb{R}^{\mathcal{A}}$. Clearly, the element-wise expression (NLC) is a special case of the more general model (56), and both are a direct consequence of (55). In fact, as we show below, the two definitions are equivalent:

Lemma 1. *Fix a propensity score profile $y \in \mathbb{R}^m$. Then, under (NLC), we have:*

1. *For all $S \in \mathcal{P}_\bullet$, the class selection probabilities $P_S(y) = \sum_{\alpha \in S} P_\alpha(y)$ satisfy (56).*
2. *For every child/parent pair $S' \triangleleft S \in \mathcal{P}_\bullet$, the conditional class selection probabilities $P_{S'|S}(y) = P_{S'}(y)/P_S(y)$ satisfy (55).*

Proof. Our proof is by reverse induction on the level $\ell = \text{lvl}(S)$ of S . For our first claim, the case $\text{lvl}(S) = L$ is simply the definition of $P_S(y)$ so there is nothing to show. Then, assuming inductively that our assertion holds for some $\ell < L$, a simple summation gives

$$\begin{aligned} P_{S_{\ell-1}}(y) &= \sum_{S_\ell \triangleleft S_{\ell-1}} P_{S_\ell}(y) = \sum_{S_\ell \triangleleft S_{\ell-1}} \prod_{k=1}^\ell \frac{\exp(y_{S_k}/\mu_k)}{\exp(y_{S_{k-1}}/\mu_k)} \\ &= \prod_{k=1}^{\ell-1} \frac{\exp(y_{S_k}/\mu_k)}{\exp(y_{S_{k-1}}/\mu_k)} \times \sum_{S_\ell \triangleleft S_{\ell-1}} \frac{\exp(y_{S_\ell}/\mu_\ell)}{\exp(y_{S_{\ell-1}}/\mu_\ell)} = \prod_{k=1}^{\ell-1} \frac{\exp(y_{S_k}/\mu_k)}{\exp(y_{S_{k-1}}/\mu_k)} \end{aligned} \quad (57)$$

where the last equality is an immediate consequence of the definition (54) of the class scores y_{S_ℓ} . This completes the induction and proves our first claim; for our second claim, simply take the quotient $P_{S'}(y)/P_S(y)$ and apply the first. ■

By Lemma 1, (NLC) and Eqs. (55) and (56) are all equivalent in terms of action distributions, so we will refer to either one as the definition of nested logit choice.

5.3. Nested exponential weights and the nested replicator dynamics. Going back to our learning considerations, combining (NLC) with the propensity scoring mechanism (51) yields the *nested exponential weights* scheme

$$\dot{y} = v(x) \quad x = P(y) \quad (\text{NEW})$$

with $P: \mathbb{R}^A \rightarrow \mathcal{X}$ defined via satisfies (NLC). In words, agents following (NEW) continue to tally the payoffs of their individual actions as they accrue over time; however, instead of disregarding intrinsic similarities between actions, they now *a*) iteratively assign a propensity score via (54) to each class of strategies that share one or more similarity attributes; and *b*) they employ the nested logit choice rule (55) to select an action based on this iterative assignment of propensities.

In the rest of this section, we provide a broad generalization of the link between (RD) and (EW) that we described in Section 5.1. Specifically, we show that, despite their completely different origins, the dynamics (NRD) and (NEW) are equivalent:

Theorem 2. *If $x(t) = P(y(t))$ is a solution orbit of (NEW), it satisfies (NRD) with revision rates*

$$\lambda_k = \begin{cases} 1/\mu_1 & \text{if } k = 0, \\ 1/\mu_{k+1} - 1/\mu_k & \text{if } k = 1, \dots, L-1. \end{cases} \quad (58)$$

Conversely, if $x(t)$ is an interior solution orbit of (NRD), it satisfies (NEW) with uncertainty parameters

$$\mu_\ell = \frac{1}{\Lambda_{\ell-1}} \equiv \frac{1}{\sum_{k=0}^{\ell-1} \lambda_k} \quad \text{for all } \ell = 1, \dots, L. \quad (59)$$

Importantly, as an immediate consequence of Theorem 2, we have:

Corollary 1. *Let $\mathcal{G} \equiv \mathcal{G}(\mathcal{A}, v)$ be a population game, and let \mathcal{P}_\bullet be a similarity structure on \mathcal{A} . Then the conclusions of Theorem 1 apply as stated to the nested dynamics (NEW).*

From a modeling viewpoint, it is worth noting that Theorem 2 reveals a somewhat unexpected relation between the revision rates λ_k of the nested protocol (NPPI) and the uncertainty parameters μ_k of the nested choice model (NLC). Mutatis mutandis, the uncertainty inherent in a given attribute partition \mathcal{P}_ℓ of \mathcal{A} is *inversely proportional* to the *total revision rate* of agents looking to switch outside the classes defined by \mathcal{P}_ℓ . Since these switches are more likely to inform the agents making them about the payoffs they could encounter by disregarding the similarities of level ℓ or lower, it is not unreasonable to expect that, in a discrete choice context, any corresponding “uncertainty” should be likewise lower when these revision rates are higher. The conversion recipe (59) can be seen as a formal version of this high-level, heuristic analogy.

The key element in the proof of Theorem 2 is Proposition 3 below, which establishes the following property of independent interest: under (NEW), the rate of change of the (recursively defined) propensity score y_S of any similarity class $S \in \mathcal{P}_\bullet$ is given by the mean class payoff $u_S(x)$ of the class in question.¹⁴ More formally, we have:

Proposition 3. *Under (NEW), we have*

$$\dot{y}_S = u_S(x) \quad \text{for all } S \in \mathcal{P}_\bullet. \quad (60)$$

¹⁴The non-nested analogue of this observation is that, under (RD)/(EW), the changes in each y_α , $\alpha \in \mathcal{A}$, under (EW) are given by the corresponding payoffs $v_\alpha(x)$ at the current mixed strategy $x = \Lambda(y)$.

Proof of Proposition 3. Our proof is by induction on the level ℓ of the class \mathcal{S}_ℓ . To begin with, for $\ell = L$, (60) is simply the definition of (NEW), so there is nothing to show. Moving forward, to establish the inductive step, assume that $\dot{y}_{\mathcal{S}_\ell} = u_{\mathcal{S}_\ell}$ for some $\ell < L$ and for all $\mathcal{S}_\ell \in \mathcal{P}_\ell$. Then, for all $\mathcal{S}_{\ell-1} \in \mathcal{P}_{\ell-1}$, we have

$$\begin{aligned} \dot{y}_{\mathcal{S}_{\ell-1}} &= \frac{d}{dt} \mu_\ell \log \sum_{\mathcal{S}_\ell \triangleleft \mathcal{S}_{\ell-1}} \exp\left(\frac{y_{\mathcal{S}_\ell}}{\mu_\ell}\right) = \mu_\ell \frac{\sum_{\mathcal{S}_\ell \triangleleft \mathcal{S}_{\ell-1}} \dot{y}_{\mathcal{S}_\ell} / \mu_\ell \cdot \exp(y_{\mathcal{S}_\ell} / \mu_\ell)}{\sum_{\mathcal{S}'_\ell \triangleleft \mathcal{S}_{\ell-1}} \exp(y_{\mathcal{S}'_\ell} / \mu_\ell)} && \# \text{ by (54)} \\ &= \sum_{\mathcal{S}_\ell \triangleleft \mathcal{S}_{\ell-1}} \dot{y}_{\mathcal{S}_\ell} \frac{x_{\mathcal{S}_\ell}}{x_{\mathcal{S}_{\ell-1}}} && \# \text{ by (55)} \\ &= \sum_{\mathcal{S}_\ell \triangleleft \mathcal{S}_{\ell-1}} x_{\mathcal{S}_\ell | \mathcal{S}_{\ell-1}} u_{\mathcal{S}_\ell}(x) && \# \text{ by induction} \\ &= \sum_{\mathcal{S}_\ell \triangleleft \mathcal{S}_{\ell-1}} \sum_{\alpha \in \mathcal{S}_\ell} x_{\alpha | \mathcal{S}_{\ell-1}} v_\alpha(x) = u_{\mathcal{S}_{\ell-1}}(x), && (61) \end{aligned}$$

where the last line following from the definition (19) of $v_{\mathcal{S}}$. This completes the induction and our proof. \blacksquare

With this structural proposition in hand, the proof of Theorem 2 proceeds as follows:

Proof of Theorem 2. Fix an alternative $\alpha \in \mathcal{A}$ with lineage $\{\alpha\} \equiv \mathcal{S}_L \triangleleft \cdots \triangleleft \mathcal{S}_1 \triangleleft \mathcal{S}_0 \equiv \mathcal{A}$. Then, by the definition of the nested logit choice rule, we have

$$\begin{aligned} x_\alpha &= \prod_{\ell=1}^L \frac{\exp(y_{\mathcal{S}_\ell} / \mu_\ell)}{\exp(y_{\mathcal{S}_{\ell-1}} / \mu_\ell)} && \# \text{ by (NLC)} \\ &= \exp\left(-\frac{y_{\mathcal{A}}}{\mu_1}\right) \times \prod_{k=1}^{L-1} \exp\left(\left(\frac{1}{\mu_k} - \frac{1}{\mu_{k+1}}\right) y_{\mathcal{S}_k}\right) \times \exp\left(\frac{y_\alpha}{\mu_L}\right) \\ &= \exp(y_\alpha / \mu_0) \times \prod_{k=0}^{L-1} \exp(-\lambda_k y_{\mathcal{S}_k}), && (62) \end{aligned}$$

with λ_k defined as in (58). Now, if we use (62) to view x_α as a function of the associated class scores $y_{\mathcal{S}_\ell}$ (viewed momentarily as independent variables), a straightforward differentiation yields

$$\frac{\partial x_\alpha}{\partial y_{\mathcal{S}_k}} = \begin{cases} -\lambda_k x_\alpha & \text{if } k \in \{0, \dots, L-1\}, \\ x_\alpha / \mu_L & \text{if } \ell = L. \end{cases} \quad (63)$$

Thus, by Proposition 3, we obtain

$$\begin{aligned} \dot{x}_\alpha &= \sum_{\ell=0}^L \frac{\partial x_\alpha}{\partial y_{\mathcal{S}_\ell}} \dot{y}_{\mathcal{S}_\ell} = x_\alpha \left[\frac{v_\alpha(x)}{\mu_L} - \sum_{k=0}^{L-1} \lambda_k u_{\mathcal{S}_k}(x) \right] \\ &= x_\alpha \left[\sum_{k=0}^{L-1} \lambda_k \cdot v_\alpha(x) - \sum_{k=0}^{L-1} \lambda_k u_{\mathcal{S}_k}(x) \right] && \# \text{ by (59)} \\ &= x_\alpha \sum_{k=0}^{L-1} \lambda_k [v_\alpha(x) - u_{\mathcal{S}_k}(x)]. && (64) \end{aligned}$$

This shows that $x_\alpha(t)$ follows the nested dynamics (NRD), as was to be shown. \blacksquare

6. LINKS WITH REGULARIZED LEARNING

We conclude our paper with a discussion of an important link between (NRD) and a more general class of game dynamics driven by reinforcement and regularization. To provide the necessary context, McFadden (1978) presents (NLC) as the choice probabilities of an additive random utility model (ARUM) whose payoff disturbances follow a generalized-extreme-value joint distribution, and shows that choice probabilities under (NLC) can be obtained as the logarithmic derivatives of the (recursively-defined) generator of this distribution.¹⁵ Expanding on previous work by Verboven (1996) and Martin et al. (2022), we provide here a representation of the choice probabilities $P_\alpha(y)$ of (NLC) and the smoothed maximal scores y_{S_ℓ} of (54) as solutions to a nonrecursive optimization problem, and we connect them to a model of regularized learning in games known as “follow the regularized leader” (FTRL).

To set the stage for all this, recall that the logit choice rule (LC) can be rewritten as

$$\Lambda(y) = \nabla y_{\mathcal{A}} = \arg \max_{x \in \mathcal{X}} \{ \langle y, x \rangle - h(x) \} \quad (65)$$

where

$$y_{\mathcal{A}} = \log \sum_{\alpha \in \mathcal{A}} \exp(y_\alpha) = \max_{x \in \mathcal{X}} \{ \langle y, x \rangle - h(x) \}. \quad (66)$$

In the above,

$$h(x) = \mu \sum_{\alpha \in \mathcal{A}} x_\alpha \log x_\alpha \quad (67)$$

denotes the negative entropy of the population state $x \in \mathcal{X}$, and we are using the standard continuity convention $0 \log 0 = 0$. Thus, given a propensity score vector $y \in \mathbb{R}^{\mathcal{A}}$, (65) states that logit choice maximizes the average score $\langle y, x \rangle$ minus an entropic penalty term; the corresponding maximum value is then provided by the softmax function (66).

From an evolutionary perspective, this implies that the dynamics (RD)/(EW) can be seen as a special case of the *regularized learning* dynamics

$$\dot{y} = v(x) \quad x = \arg \max_{x' \in \mathcal{X}} \{ \langle y, x' \rangle - h(x') \} \quad (\text{RL})$$

where $h: \mathcal{X} \rightarrow \mathbb{R}$ is a generic penalty term, referred to as the method’s “regularizer” (Hofbauer & Sandholm, 2009; Shalev-Shwartz, 2011; Mertikopoulos & Sandholm, 2016). In words, (RL) admits a straightforward interpretation: agents simply adjust their strategies over time by best-responding to their actions’ cumulative payoffs minus a regularization penalty intended to incentivize exploration. This regularization idea is essentially the one underlying the stochastic – or “smooth”, depending on the context – fictitious play process of Fudenberg & Kreps (1993), which has been studied in a wide variety of contexts by Fudenberg & Levine (1995, 1999), Hofbauer & Sandholm (2009), Bravo & Mertikopoulos (2017), Giannou et al. (2021), and many others; for a partial survey, see Hofbauer & Sandholm (2009) and references therein.¹⁶

In the remaining of this section, we extend this interpretation to the nested dynamics (NRD)/(NEW). This extension is quite intricate and hinges on a delicately tuned, nested variant of the entropic regularizer (67) defined as

$$h(x) = \sum_{k=0}^L \delta_k \sum_{S_k \in \mathcal{P}_k} x_{S_k} \log x_{S_k} \quad (68)$$

¹⁵For its part, the logit rule (LC) requires extreme-value distributed payoff disturbances that are independent across strategies; for more details, see Anderson et al. (1992).

¹⁶In the online learning literature, (RL) is often referred to as “follow the regularized leader” (FTRL), cf. Shalev-Shwartz & Singer (2006) and Shalev-Shwartz (2011).

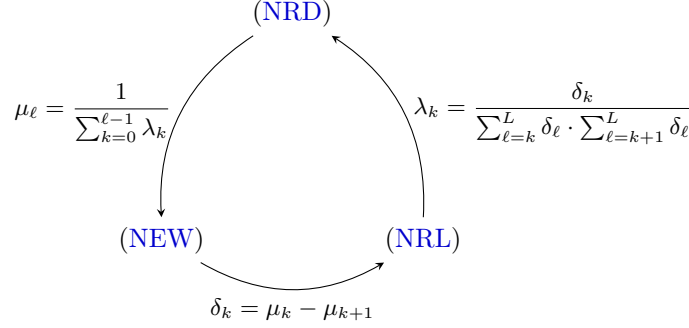


Figure 5: The triple equivalence between (NRD), (NEW), and (NRL).

where the weight coefficients δ_k , $k = 0, \dots, L$, are given by

$$\delta_k = \begin{cases} \mu_k - \mu_{k+1} & \text{for } k = 0, \dots, L-1, \\ \mu_L & \text{for } k = L. \end{cases} \quad (69)$$

A version of this “nested entropy” was first derived by Mertikopoulos & Sandholm (2018) and Martin et al. (2022) in the context of Riemannian game dynamics and multi-armed bandits respectively. In the present setting, an interpretation of this definition is as follows: at the finest level ($k = L$), the term $\sum_{\alpha \in \mathcal{A}} x_\alpha \log x_\alpha$ measures the (negative) entropy of the population state x , with no similarities taken into account. Then, at each progressively coarser level $k = L-1, \dots, 0$, the various summands of the form $\sum_{S_k \in \mathcal{P}_k} x_{S_k} \log x_{S_k}$ measure the entropy of the probability distribution induced by x on $\Delta(\mathcal{P}_k)$, that is, restricted to k -tier similarity classes. These individual entropy measures do not enter the definition of h_S with uniform weight, but are suitably adjusted by the uncertainty level μ_ℓ associated to each similarity partition \mathcal{P}_ℓ of \mathcal{P}_\bullet ; the specific choice of the weight coefficients involves a significant degree of hindsight, but the fact that it completes the triangle with (36) and (58) is, of course, not a coincidence.

With this definition in hand, we have the following equivalence:

Theorem 3. *For all $y \in \mathbb{R}^{\mathcal{A}}$, the probability distribution (NLC) can be written as*

$$P(y) = \arg \max_{x \in \mathcal{X}} \{ \langle y, x \rangle - h(x) \} \quad (70)$$

with $h(x)$ given by (68). Consequently $x(t) = P(y(t))$ is a solution orbit of (NEW) if and only if it is a solution orbit of the nested regularized learning dynamics

$$\dot{y} = v(x) \quad x = \arg \max_{x' \in \mathcal{X}} \left\{ \langle y, x' \rangle - \sum_{k=0}^L \delta_k \sum_{S_k \in \mathcal{P}_k} x'_{S_k} \log x'_{S_k} \right\} \quad (\text{NRL})$$

with δ_k given by (69).

Thus, combining Theorems 2 and 3, we obtain a triple equivalence between the revisionist, reinforcement, and regularized learning viewpoints, as represented schematically in Fig. 5 (the proof of Theorem 3 is presented in Appendix B). Importantly, the regularized dynamics (RL) – of which (NRL) constitute a special case – have been studied extensively in the online learning literature, so this triple equivalence can be used to infer a number of additional properties for (NRD) which would otherwise be particularly difficult to piece together – from recurrence in zero-sum (Mertikopoulos et al., 2018) and harmonic games (Legacci et al.,

2024) to evolutionary permanence and impermanence in the spirit of Hofbauer & Sigmund (1988). We defer such examinations to future work.

APPENDIX A. PROOF OF THEOREM 1

Our aim in this appendix is to prove Theorem 1 and Proposition 1 on the rationality and convergence properties of the dynamics (NRD)/(NEW). We begin with a short calculation which is very useful in several parts of our proof:

Lemma A.1. *Let $x(t)$ be a solution orbit of (NRD), and set*

$$W_{\alpha\beta}(t) = D_L(e_\alpha, x(t)) - D_L(e_\beta, x(t)) = \sum_{\ell=1}^L \delta_\ell \log \frac{x_{[\beta]_\ell}(t)}{x_{[\alpha]_\ell}(t)}, \quad (\text{A.1})$$

for all $\alpha, \beta \in \mathcal{A}$. Then $\dot{W}_{\alpha\beta}(t) = v_\beta(x(t)) - v_\alpha(x(t))$.

Proof. By Proposition 2, we readily get

$$\dot{W}_{\alpha\beta} = \langle v(x), x - e_\alpha \rangle - \langle v(x), x - e_\beta \rangle = \langle v(x), e_\beta - e_\alpha \rangle = v_\beta(x) - v_\alpha(x). \quad \blacksquare$$

We are now in a position to carry out the proof of Theorem 1, one property at a time.

Proof of Part 1: Elimination of strictly dominated strategies. Suppose that $\alpha \in \mathcal{A}$ is strictly dominated by $\beta \in \mathcal{A}$ and let $x(t)$ be a solution of (NRD) with $x(0) \in \text{ri}(\mathcal{X})$, so $W_{\alpha\beta}(0) < \infty$. Then, Lemma A.1 implies that $\dot{W}_{\alpha\beta} \geq c$ for some $c > 0$, from which it follows that

$$-\sum_{\ell=1}^L \delta_\ell \log x_{[\alpha]_\ell}(t) \geq W_{\alpha\beta}(t) \geq W_{\alpha\beta}(0) + ct \rightarrow \infty \quad \text{as } t \rightarrow \infty. \quad (\text{A.2})$$

Since $x_{[\alpha]_\ell} \geq x_\alpha$ for all $\ell = 1, \dots, L$, we obtain

$$\sum_{\ell=1}^L \delta_\ell \log x_{[\alpha]_\ell}(t) \geq \sum_{\ell=1}^L \delta_\ell \log x_\alpha(t) \quad (\text{A.3})$$

so, by (A.2), we conclude that $x_\alpha(t) \rightarrow 0$ as $t \rightarrow \infty$, i.e., α becomes extinct along $x(t)$. \blacksquare

Proof of Part 2: Stationary points. By definition, x^* is a restricted equilibrium of \mathcal{G} if and only if $v_\alpha(x^*) = v_\beta(x^*)$ whenever $\alpha, \beta \in \text{supp}(x^*)$. Thus, stationarity of restricted equilibria in (NRD) follows immediately by observing that $v_\alpha(x^*) = u_{[\alpha]_\ell}(x^*)$ whenever $x_\alpha^* > 0$.

Conversely, assume that x^* is a stationary point of (NRD) which is not a restricted equilibrium of \mathcal{G} . Then, for all $\alpha \in \arg \max_\beta v_\beta(x^*)$ and all $\ell = 1, \dots, L$, we have $v_\alpha(x^*) \geq u_{[\alpha]_\ell}(x^*)$, with the inequality being strict for $\ell < L$. Since $\lambda_0 > 0$, it follows that the right-hand side of (NRD) is strictly positive at x^* , contradicting the assumption that x^* is stationary. \blacksquare

Proof of Parts 3 and 4: Limits of interior solutions and Lyapunov stability. We will prove both claims in tandem by showing the following more general result: if every neighborhood \mathcal{U} of $x^* \in \mathcal{X}$ admits an interior solution $x(t)$ that remains in \mathcal{U} for all $t \geq 0$, then x^* is a Nash equilibrium.

To establish this more general claim, assume to the contrary that x^* is not a Nash equilibrium, so $v_\alpha(x^*) < v_\beta(x^*)$ for some $\alpha \in \text{supp}(x^*)$, $\beta \in \mathcal{A}$. Then there exists a neighborhood \mathcal{U} of x^* in \mathcal{X} such that $x_\alpha > x_\alpha^*/2$ and $v_\beta(x) - v_\alpha(x) \geq c$ for some $c > 0$ and for all $x \in \mathcal{U}$. By assumption, there exists an interior solution orbit $x(t)$ of (NRD) that is contained in \mathcal{U} for all $t \geq 0$. Since $v_\alpha(x(t)) - v_\beta(x(t)) < -c$, invoking Lemma A.1 in the

same way as in the proof of Part 1 implies that $W_{\alpha\beta}(t) \rightarrow \infty$ as $t \rightarrow \infty$. However, since $x_\alpha(t) > x_\alpha^*/2$ for all $t \geq 0$ by the definition of \mathcal{U} , we also have that

$$W_{\alpha\beta}(t) \leq \sum_{\ell=1}^L \log \frac{x_{[\beta]\ell}(t)}{x_\alpha^*/2} \leq \sum_{\ell=1}^L \log \frac{2}{x_\alpha^*} < \infty, \quad (\text{A.4})$$

a contradiction (recall here that $\alpha \in \text{supp}(x^*)$). In turn, going back to our original working assumption, we conclude that x^* must be a Nash equilibrium of \mathcal{G} , as claimed. ■

Proof of Part 5: Convergence to evolutionarily stable states. We consider the case in which x^* is a GESS; our claim for non-global ESSs follows by a straightforward localization of our argument to a sufficiently small neighborhood of x^* defined by a level set of $D_L(x^*, \cdot)$.

To begin, define the *effective domain* (x^*) of $D_L(x^*, \cdot)$ as

$$(x^*) \equiv \{x \in \mathcal{X} : D_L(x^*, x) < \infty\} = \{x \in \mathcal{X} : \text{supp}(x^*) \subseteq \text{supp}(x)\}. \quad (\text{A.5})$$

Since every solution of (NRD) has constant support, (x^*) is an invariant set of (NRD). Moreover, for any solution $x(t)$ contained in (x^*) , Proposition 2 and the definition of GESS imply that

$$\frac{d}{dt} D_L(x^*, x(t)) = \langle v(x(t)), x(t) - x^* \rangle \leq 0, \quad (\text{A.6})$$

with equality if and only if $x(t) = x^*$.¹⁷

Consequently, to establish that x^* is asymptotically stable with basin (x^*) , it suffices to show that $x(t)$ has no ω -limit points in $\mathcal{X} \setminus (x^*)$ (cf. Sandholm, 2010a, Appendix 7.B). To do so, assume to the contrary that $x(t)$ admits an ω -limit $\hat{x} \neq x^*$, so $x(t_n) \rightarrow \hat{x}$ for some divergent sequence of times $t_n \uparrow \infty$. By the definition of (NRD), we further have

$$\begin{aligned} |\dot{x}_\alpha| &\leq x_\alpha \sum_{k=0}^{L-1} \lambda_k |v_\alpha(x) - v_{[\alpha]_k}(x)| \\ &\leq \sum_{k=0}^{L-1} \lambda_k \cdot \max_{x \in \mathcal{X}} |v_\alpha(x) - v_{[\alpha]_k}(x)| \leq 2 \max_{x \in \mathcal{X}} \|v(x)\|_\infty \end{aligned} \quad (\text{A.7})$$

where, in the last line, we used the facts that $x_\alpha \leq 1$ and that $|u_S(x)| \leq \max_{\alpha \in \mathcal{S}} |v_\alpha(x)|$.¹⁸ It thus follows that there exists an open neighborhood \mathcal{U} of \hat{x} , scalars $c, \delta > 0$ and an index $n_0 \geq 1$ such that $x(t) \in \mathcal{U}$ and $\langle v(x(t)), x(t) - x^* \rangle \leq -c < 0$ for all $t \in [t_n, t_n + \delta]$ and all $n \geq n_0$. Hence, using (A.6), we get

$$D_L(x^*, x(t_n + \delta)) - D_L(x^*, x(0)) \leq \int_0^{t_n + \delta} \langle v(x(s)), x(s) - x^* \rangle d\tau \leq -c\delta(n - n_0). \quad (\text{A.8})$$

Hence, letting $n \rightarrow \infty$, we get $\liminf_{t \rightarrow \infty} D_L(x^*, x(t)) = -\infty$, a contradiction which proves our original assertion. ■

Proof of Part 6: Convergence in potential games. We will establish the dynamics' convergence in potential games by showing that the game's potential is a strict (increasing) Lyapunov function for (NRD). Indeed, if $v = \nabla F$ for some potential function F on \mathcal{X} , we have:

$$\dot{F} = \sum_{\alpha \in \mathcal{A}} \frac{\partial F}{\partial x_\alpha} \dot{x}_\alpha = \sum_{\alpha \in \mathcal{A}} v_\alpha(x) x_\alpha \sum_{k=0}^{L-1} \lambda_k [v_\alpha(x) - u_{[\alpha]_k}(x)] \quad \# \text{ by (NRD)}$$

¹⁷In other words, $D_L(x^*, x(t))$ is a strict Lyapunov function for (NRD) on (x^*) .

¹⁸The notation $\|\cdot\|_\infty$ stands here for the sup-norm on \mathbb{R}^A , viz. $\|\pi\|_\infty = \max_\alpha |\pi_\alpha|$.

$$\begin{aligned}
&= \sum_{k=0}^{L-1} \lambda_k \sum_{\alpha \in \mathcal{A}} x_\alpha v_\alpha(x) [v_\alpha(x) - u_{[\alpha]_k}(x)] \quad \# \text{ rearrange product} \\
&= \sum_{k=0}^{L-1} \lambda_k \sum_{\mathcal{S}_k \in \mathcal{P}_k} \sum_{\alpha \in \mathcal{S}_k} x_\alpha v_\alpha(x) [v_\alpha(x) - u_{\mathcal{S}_k}(x)] \quad \# \text{ collect parent classes} \\
&= \sum_{k=0}^{L-1} \lambda_k \sum_{\mathcal{S}_k \in \mathcal{P}_k} x_{\mathcal{S}_k} \left[\sum_{\alpha \in \mathcal{S}_k} x_{\alpha|\mathcal{S}_k} v_\alpha^2(x) - \left[\sum_{\alpha \in \mathcal{S}_k} x_{\alpha|\mathcal{S}_k} v_\alpha(x) \right]^2 \right]. \quad (\text{A.9})
\end{aligned}$$

By Jensen's inequality applied to the vector of conditional probabilities $(x_{\alpha|\mathcal{S}_k})_{\alpha \in \mathcal{S}_k} \in \Delta(\mathcal{S}_k)$, we conclude that every inner summand in (A.9) is non-negative and, moreover, it vanishes if and only if $x_{\alpha|\mathcal{S}_k} = 1$ for some $\alpha \in \mathcal{S}_k$.¹⁹ We thus conclude that $\dot{F}(x(t)) \geq 0$ with equality if and only if $x(t)$ is a restricted equilibrium of \mathcal{G} . Our claim then follows from Part 2 of our theorem combined with standard Lyapunov function arguments – cf. Sandholm (2010a, Appendix 7.B). ■

Proof of Part 7: Convergence in strictly monotone games. Let x^* be a Nash equilibrium of \mathcal{G} (that every population game $\mathcal{G} \equiv \mathcal{G}(\mathcal{A}, v)$ admits an equilibrium follows from standard fixed point arguments, cf. Sandholm, 2010a). Then, by combining the variational characterization (7) of x^* and the monotonicity postulate (9) for \mathcal{G} , we readily get

$$\langle v(x), x - x^* \rangle \leq \langle v(x^*), x - x^* \rangle \leq 0 \quad (\text{A.10})$$

with equality holding if and only $x = x^*$ (by the strict monotonicity of \mathcal{G}). In turn, this implies that the (necessarily unique) equilibrium x^* of a strictly monotone game is a GESS thereof, so our convergence claim follows from Part 5 of the theorem. ■

We conclude this appendix with the proof of our result on the rate of extinction of dominated strategies in the presence of similarities.

Proof of Proposition 1. Let $\ell = \deg(\alpha, \beta)$ be the degree of similarity between α and β . Then, by the definition (A.1) of $W_{\alpha\beta}$, we have

$$W_{\alpha\beta}(t) = \sum_{k=1}^L \delta_k \log \frac{x_{[\beta]_k}(t)}{x_{[\alpha]_k}(t)} = \sum_{k=\ell+1}^L \delta_k \log \frac{x_{[\beta]_k}(t)}{x_{[\alpha]_k}(t)} \quad (\text{A.11})$$

on account of the fact that $[\alpha]_k = [\beta]_k$ for all $k = 1, \dots, \ell$. Since $x_{[\beta]_k} \leq 1$ and $x_{[\alpha]_k} \geq x_\alpha$ for all $k = \ell + 1, \dots, L$, we further get

$$- \sum_{k=\ell+1}^L \delta_k \log x_\alpha(t) \geq \sum_{k=\ell+1}^L \delta_k \log \frac{x_{[\beta]_k}(t)}{x_{[\alpha]_k}(t)} = W_{\alpha\beta}(t) \geq W_{\alpha\beta}(0) + ct \quad (\text{A.12})$$

by (A.2). However, by (44), we have

$$\sum_{k=\ell+1}^L \delta_k = \frac{1}{\Lambda_\ell} = \frac{1}{\sum_{k=0}^\ell \lambda_k} \quad (\text{A.13})$$

so our assertion follows by substituting the above in (A.12) and solving for $x_\alpha(t)$. ■

¹⁹Note also that each of these terms can be interpreted as the conditional variance of $v_\alpha(x)$ restricted to class \mathcal{S}_k of \mathcal{P}_k .

APPENDIX B. PROOF OF THEOREM 3

Our aim in this appendix is to prove the representation of the nested logit choice rule (NLC) as a regularized best response model (cf. Theorem 3 in Section 5.2).²⁰ To that end, given a similarity structure \mathcal{P}_\bullet on \mathcal{A} and a sequence of uncertainty parameters $\mu_1 \geq \dots \geq \mu_L > 0$ (with $\mu_{L+1} = 0$ by convention), it will be convenient to introduce the following (conditional) measures of entropy:

- (1) The *nested entropy* of $x \in \Delta(\mathcal{A})$ relative to $\mathcal{S} \in \mathcal{P}_\ell$:

$$h_{\mathcal{S}}(x) = \sum_{k=\ell}^L \delta_k \sum_{\mathcal{S}_k \preccurlyeq_k \mathcal{S}} x_{\mathcal{S}_k} \log x_{\mathcal{S}_k} \quad (\text{B.1})$$

where $\delta_k = \mu_k - \mu_{k+1}$ for all $k = 1, \dots, L$.

- (2) The *restricted entropy* of $x \in \Delta(\mathcal{A})$ relative to $\mathcal{S} \in \mathcal{P}_\ell$:

$$h_{|\mathcal{S}}(x) = h_{\mathcal{S}}(x) + \chi_{\Delta(\mathcal{S})}(x) = \begin{cases} h_{\mathcal{S}}(x) & \text{if } x \in \Delta(\mathcal{S}), \\ \infty & \text{otherwise,} \end{cases} \quad (\text{B.2})$$

where $\chi_{\Delta(\mathcal{S})}$ denotes the characteristic function of $\Delta(\mathcal{S})$, viz. $\chi_{\Delta(\mathcal{S})}(x) = 0$ if $x \in \Delta(\mathcal{S})$ and $\chi_{\Delta(\mathcal{S})}(x) = \infty$ otherwise. [Obviously, $h_{|\mathcal{S}}(x) = h_{\mathcal{S}}(x)$ whenever $x \in \Delta(\mathcal{S})$.]

- (3) The *conditional entropy* of $x \in \Delta(\mathcal{A})$ relative to a target class $\mathcal{S} \in \mathcal{P}_\ell$:

$$h(x|\mathcal{S}) = \mu_{\ell+1} \sum_{\mathcal{S}' \triangleleft \mathcal{S}} x_{\mathcal{S}'} \log \frac{x_{\mathcal{S}'}}{x_{\mathcal{S}}} = \mu_{\ell+1} x_{\mathcal{S}} \sum_{\mathcal{S}' \triangleleft \mathcal{S}} x_{\mathcal{S}'|\mathcal{S}} \log x_{\mathcal{S}'|\mathcal{S}}. \quad (\text{B.3})$$

Remark 1. As per our standard conventions, we are treating \mathcal{S} interchangeably as a subset of \mathcal{A} or as an element of \mathcal{P}_\bullet ; by analogy, to avoid notational inflation, we are also viewing $\Delta(\mathcal{S})$ as a subset of $\Delta(\mathcal{A})$ – more precisely, a face thereof. Finally, in all cases, the functions $h(x|\mathcal{S})$, $h_{\mathcal{S}}(x)$ and $h_{|\mathcal{S}}(x)$ are assumed to take the value $+\infty$ for $x \in \mathbb{R}^{\mathcal{A}} \setminus \Delta(\mathcal{A})$. §

Remark 2. For posterity, we also note that the nested and restricted entropy functions ($h_{\mathcal{S}}(x)$ and $h_{|\mathcal{S}}(x)$ respectively) are both convex over $\Delta(\mathcal{A})$. This is a consequence of the fact that each summand $x_{\mathcal{S}} \log x_{\mathcal{S}}$ in (B.1) is convex in x and that $\delta_k = \mu_k - \mu_{k+1} \geq 0$ for all $k = 1, \dots, L$. Of course, any two distributions $x, x' \in \Delta(\mathcal{A})$ that assign the same probabilities to elements of \mathcal{S} but not otherwise have $h_{\mathcal{S}}(x) = h_{\mathcal{S}}(x')$, so $h_{\mathcal{S}}$ is *not* strictly convex over $\Delta(\mathcal{A})$ if $\mathcal{S} \neq \mathcal{A}$. However, since the function $\sum_{\alpha \in \mathcal{S}} x_{\alpha} \log x_{\alpha}$ is strictly convex over $\Delta(\mathcal{S})$, it follows that $h_{\mathcal{S}}$ – and hence $h_{|\mathcal{S}}$ – is strictly convex over $\Delta(\mathcal{S})$. §

As a preamble to proving Theorem 3, we will require the technical result below, which illustrates the precision connection between the nested and conditional entropy functions:

Proposition B.1. *For all $\mathcal{S} \in \mathcal{P}_\ell$, $\ell = 1, \dots, L$, and for all $x \in \Delta(\mathcal{A})$, we have:*

$$h_{\mathcal{S}}(x) = \sum_{\mathcal{S}' \preccurlyeq \mathcal{S}} h(x|\mathcal{S}') + \mu_{\ell} x_{\mathcal{S}} \log x_{\mathcal{S}}. \quad (\text{B.4})$$

Consequently, for all $x \in \Delta(\mathcal{S})$, we have:

$$h_{|\mathcal{S}}(x) = \sum_{\mathcal{S}' \preccurlyeq \mathcal{S}} h(x|\mathcal{S}'). \quad (\text{B.5})$$

²⁰The proof presented here closely follows the conference paper of Martin et al. (2022), which was published when the current paper was under review.

Proof. Let $\ell = \text{lvl}(\mathcal{S})$, and fix some attribute label $k > \ell$. We will proceed inductively by collecting all terms in (B.4) associated to the attribute \mathcal{P}_k and then summing everything together. Indeed, we have:

$$\begin{aligned} \mu_k \sum_{S' \preccurlyeq_k \mathcal{S}} x_{S'} \log x_{S'} &= \mu_k \sum_{\mathcal{S}_{k-1} \preccurlyeq_{k-1} \mathcal{S}} \left[\sum_{S' \triangleleft \mathcal{S}_{k-1}} x_{S'} \log x_{S'} \right] && \# \text{ collect attributes} \\ &= \mu_k \sum_{\mathcal{S}_{k-1} \preccurlyeq_{k-1} \mathcal{S}} \left[\sum_{S' \triangleleft \mathcal{S}_{k-1}} x_{S'|\mathcal{S}_{k-1}} x_{\mathcal{S}_{k-1}} \log(x_{S'|\mathcal{S}_{k-1}} x_{\mathcal{S}_{k-1}}) \right] \\ &= \mu_k \sum_{\mathcal{S}_{k-1} \preccurlyeq_{k-1} \mathcal{S}} \left[\sum_{S' \triangleleft \mathcal{S}_{k-1}} x_{S'|\mathcal{S}_{k-1}} x_{\mathcal{S}_{k-1}} \log x_{S'|\mathcal{S}_{k-1}} \right] && \text{(B.6a)} \end{aligned}$$

$$+ \mu_k \sum_{\mathcal{S}_{k-1} \preccurlyeq_{k-1} \mathcal{S}} \left[\sum_{S' \triangleleft \mathcal{S}_{k-1}} x_{S'|\mathcal{S}_{k-1}} x_{\mathcal{S}_{k-1}} \log x_{\mathcal{S}_{k-1}} \right] \quad \text{(B.6b)}$$

with the tacit understanding that any empty sum that appears above is taken equal to zero.

Now, by the definition of the nested entropy, we readily obtain that

$$\text{(B.6a)} = \sum_{\mathcal{S}_{k-1} \preccurlyeq_{k-1} \mathcal{S}} h(x|\mathcal{S}_{k-1}) \quad \text{(B.7a)}$$

whereas, by noting that $\sum_{S' \triangleleft \mathcal{S}_{k-1}} x_{S'|\mathcal{S}_{k-1}} = 1$ (by the definition of conditional class choice probabilities), Eq. (B.6b) becomes

$$\text{(B.6b)} = \mu_k \sum_{\mathcal{S}_{k-1} \preccurlyeq_{k-1} \mathcal{S}} x_{\mathcal{S}_{k-1}} \log x_{\mathcal{S}_{k-1}}. \quad \text{(B.7b)}$$

Hence, combining Eqs. (B.6), (B.7a) and (B.7b), we get:

$$\mu_k \sum_{S' \preccurlyeq_k \mathcal{S}} x_{S'} \log x_{S'} = \sum_{\mathcal{S}_{k-1} \preccurlyeq_{k-1} \mathcal{S}} h(x|\mathcal{S}_{k-1}) + \mu_k \sum_{\mathcal{S}_{k-1} \preccurlyeq_{k-1} \mathcal{S}} x_{\mathcal{S}_{k-1}} \log x_{\mathcal{S}_{k-1}}. \quad \text{(B.8)}$$

The above expression is the basic building block of our inductive construction. Indeed, summing (B.8) over all $k = L, \dots, \ell = \text{lvl}(\mathcal{S})$, we obtain:

$$\begin{aligned} h_{\mathcal{S}}(x) &= \sum_{k=\ell}^L (\mu_k - \mu_{k+1}) \sum_{S' \preccurlyeq_k \mathcal{S}} x_{S'} \log x_{S'} && \# \text{ by definition} \\ &= \sum_{k=L}^{\ell+1} \left[\mu_k \sum_{S' \preccurlyeq_k \mathcal{S}} x_{S'} \log x_{S'} - \mu_{k+1} \sum_{S' \preccurlyeq_k \mathcal{S}} x_{S'} \log x_{S'} \right] + (\mu_{\ell} - \mu_{\ell+1}) x_{\mathcal{S}} \log x_{\mathcal{S}} \\ &&& \# \text{ isolate } \mathcal{S} \\ &= \sum_{k=L}^{\ell+1} \left[\sum_{\mathcal{S}_{k-1} \preccurlyeq_{k-1} \mathcal{S}} h(x|\mathcal{S}_{k-1}) + \mu_k \sum_{\mathcal{S}_{k-1} \preccurlyeq_{k-1} \mathcal{S}} x_{\mathcal{S}_{k-1}} \log x_{\mathcal{S}_{k-1}} - \mu_{k+1} \sum_{S' \preccurlyeq_k \mathcal{S}} x_{S'} \log x_{S'} \right] \\ &&& \# \text{ by (B.8)} \\ &= \sum_{k=\ell}^{L-1} \sum_{S' \preccurlyeq_k \mathcal{S}} h(x|S') + \mu_{\ell} x_{\mathcal{S}} \log x_{\mathcal{S}} - \mu_{L+1} \sum_{S' \preccurlyeq_L \mathcal{S}} x_{S'} \log x_{S'} \quad \text{(B.9)} \end{aligned}$$

with the last equality following by telescoping the terms involving μ_k . Now, given that $\mu_{L+1} = 0$ by convention, the third sum above is zero. Finally, since the conditional entropy

of x relative to any childless class is zero by definition, the first sum in (B.9) can be rewritten as $\sum_{k=\ell}^{L-1} \sum_{S' \preccurlyeq_k S} h(x|S') = \sum_{S' \preccurlyeq S} h(x|S')$, and our claim follows.

Finally, (B.5) is a consequence of the fact that $x_S = 1$ whenever $x \in \Delta(S)$ – i.e., whenever $\text{supp}(x) \subseteq S$. ■

We are now in a position to state and prove a class-based version of Theorem 3, which is of independent interest, and which contains Theorem 3 as a special case:

Proposition B.2. *The following holds for all $S \in \mathcal{P}_\bullet$ and all $y \in \mathbb{R}^A$:*

1. *The recursively defined propensity score y_S of S as given by (54) can be expressed as*

$$y_S = \max_{x \in \Delta(S)} \{\langle y, x \rangle - h_S(x)\} \quad (\text{B.10})$$

2. *The conditional probability of choosing $\alpha \in \mathcal{A}$ given that S has already been selected under (55) is given by*

$$P_{\alpha|S}(y) = \frac{\partial y_S}{\partial y_\alpha} \quad (\text{B.11})$$

and the conditional probability vector $P_{|S}(y) = (P_{\alpha|S}(y))_{\alpha \in \mathcal{A}}$ solves the problem (B.10), viz.

$$P_{|S}(y) = \arg \max_{x \in \Delta(S)} \{\langle y, x \rangle - h_S(x)\} \quad (\text{B.12})$$

Proof. We begin by noting that the optimization problem (B.10) can be written more explicitly as

$$\begin{aligned} & \text{maximize} \quad \langle y, x \rangle - h_S(x), \\ & \text{subject to} \quad x \in \Delta(\mathcal{A}) \text{ and } \text{supp}(x) \subseteq S. \end{aligned} \quad (\text{Opt}_S)$$

We will proceed to show that the (unique) solution of (Opt_S) is given by the vector of conditional probabilities $(P_{\alpha|S}(y))_{\alpha \in \mathcal{A}}$. The expression (B.10) for the maximal value of (Opt_S) will then be derived from Proposition B.1, and the differential representation (B.11) will follow from Legendre's identity. We make all this precise in a series of individual steps below.

Step 1: Optimality conditions for (Opt_S). Fix an alternative $\alpha \in S$ with lineage $\{\alpha\} \equiv S_L \triangleleft \dots \triangleleft S_\ell \equiv S$. Then the definition of the nested entropy gives

$$\begin{aligned} \frac{\partial h_S}{\partial x_\alpha} &= \sum_{k=\ell}^L \delta_k \sum_{S' \preccurlyeq_k S} \frac{\partial}{\partial x_\alpha} (x_{S'} \log x_{S'}) = \sum_{k=\ell}^L \delta_k \sum_{S' \preccurlyeq_k S} (1 + \log x_{S'}) \frac{\partial x_{S'}}{\partial x_\alpha} \\ &= \sum_{k=\ell}^L \delta_k \sum_{S' \preccurlyeq_k S} (1 + \log x_{S'}) \mathbb{1}\{\alpha \in S'\} \\ &= \sum_{k=\ell}^L \delta_k (1 + \log x_{S_k}) \\ &= \mu_\ell + \sum_{k=\ell}^L \delta_k \log x_{S_k} \end{aligned} \quad (\text{B.13})$$

This implies that $\partial_\alpha h_S(x) \rightarrow -\infty$ whenever $x_\alpha \rightarrow 0$, so any solution x of (Opt_S) must have $x_\alpha > 0$ for all $\alpha \in S$. In view of this, the first-order optimality conditions for (Opt_S) become

$$y_\alpha - \frac{\partial h_S}{\partial x_\alpha} = y_\alpha - \mu_\ell - \sum_{k=\ell}^L \delta_k \log x_{S_k} = \lambda \quad \text{for all } \alpha \in S, \quad (\text{B.14})$$

where λ is the Lagrange multiplier for the equality constraint $\sum_{\alpha \in \mathcal{A}} x_\alpha = 1$.²¹ Thus, after rearranging terms and exponentiating, we get

$$x_{\mathcal{S}_L}^{\delta_L} \cdot x_{\mathcal{S}_{L-1}}^{\delta_{L-1}} \cdots x_{\mathcal{S}_\ell}^{\delta_\ell} = \frac{\exp(y_\alpha)}{Z}, \quad (\text{B.15})$$

for some proportionality constant $Z \equiv Z(y) > 0$.

Step 2: Solving (Opt_S). The next step of our proof will focus on unrolling the chain (B.15), one attribute at a time. To start, recall that $\delta_L = \mu_L$, so (B.15) becomes

$$x_{\mathcal{S}_L} \cdot x_{\mathcal{S}_{L-1}}^{\delta_{L-1}/\mu_L} \cdots x_{\mathcal{S}_\ell}^{\delta_\ell/\mu_L} = \frac{\exp(y_{\mathcal{S}_L}/\mu_L)}{Z^{1/\mu_L}}, \quad (\text{B.16})$$

where we used the fact that $\mathcal{S}_L = \alpha$ by definition. Now, since $\mathcal{S}_{L-1} \preceq \mathcal{S}_\ell = \mathcal{S}$, it follows that all children of \mathcal{S}_{L-1} are also descendants of \mathcal{S} , so (B.16) applies to all siblings of \mathcal{S}_L as well. Hence, summing (B.16) over $\mathcal{S}_L \triangleleft \mathcal{S}_{L-1}$, we get

$$x_{\mathcal{S}_{L-1}} \cdot x_{\mathcal{S}_{L-1}}^{\delta_{L-1}/\mu_L} \cdots x_{\mathcal{S}_\ell}^{\delta_\ell/\mu_L} = \frac{\exp(y_{\mathcal{S}_{L-1}}/\mu_L)}{Z^{1/\mu_L}}, \quad (\text{B.17})$$

where we used the definition (2) of $x_{\mathcal{S}_{L-1}} = \sum_{\mathcal{S}_L \triangleleft \mathcal{S}_{L-1}} x_{\mathcal{S}_L}$ and the recursive definition (54) for $y_{\mathcal{S}_{L-1}}$, i.e., the fact that $\exp(y_{\mathcal{S}_{L-1}}/\mu_L) = \sum_{\mathcal{S}_L \triangleleft \mathcal{S}_{L-1}} \exp(y_{\mathcal{S}_L}/\mu_L)$. Therefore, noting that

$$1 + \frac{\delta_{L-1}}{\mu_L} = 1 + \frac{\mu_{L-1} - \mu_L}{\mu_L} = \frac{\mu_{L-1}}{\mu_L} \quad (\text{B.18})$$

the product (B.17) becomes

$$x_{\mathcal{S}_{L-1}}^{\mu_{L-1}} \cdot x_{\mathcal{S}_{L-2}}^{\delta_{L-2}} \cdots x_{\mathcal{S}_\ell}^{\delta_\ell} = \frac{\exp(y_{\mathcal{S}_{L-1}})}{Z} \quad (\text{B.19})$$

or, equivalently

$$x_{\mathcal{S}_{L-1}} \cdot x_{\mathcal{S}_{L-2}}^{\delta_{L-2}/\mu_{L-1}} \cdots x_{\mathcal{S}_\ell}^{\delta_\ell/\mu_{L-1}} = \frac{\exp(y_{\mathcal{S}_{L-1}}/\mu_{L-1})}{Z^{1/\mu_{L-1}}}. \quad (\text{B.20})$$

This last equation has the same form as (B.17) applied to the chain $\mathcal{S}_\ell \triangleright \mathcal{S}_{\ell+1} \triangleright \cdots \triangleright \mathcal{S}_{L-1}$ instead of $\mathcal{S}_\ell \triangleright \mathcal{S}_{\ell+1} \triangleright \cdots \triangleright \mathcal{S}_L$. Thus, proceeding inductively, we conclude that

$$x_{\mathcal{S}_k}^{\mu_k} \prod_{j=k-1}^{\ell} x_{\mathcal{S}_j}^{\delta_j} = \frac{\exp(y_{\mathcal{S}_k})}{Z} \quad \text{for all } k = L, \dots, \ell \quad (\text{B.21})$$

with the empty product $\prod_{j \in \emptyset} x_{\mathcal{S}_j}^{\delta_j}$ taken equal to 1 by standard convention.

Now, substituting $k \leftarrow k+1$ in (B.21), we readily get

$$x_{\mathcal{S}_{k+1}}^{\mu_{k+1}} \cdot x_{\mathcal{S}_k}^{\delta_k} \prod_{j=k-1}^{\ell} x_{\mathcal{S}_j}^{\delta_j} = \frac{\exp(y_{\mathcal{S}_{k+1}})}{Z} \quad \text{for all } k = L-1, \dots, \ell. \quad (\text{B.22})$$

Consequently, recalling that $\delta_k = \mu_k - \mu_{k+1}$ and dividing (B.21) by (B.22), we get

$$\frac{x_{\mathcal{S}_{k+1}}^{\mu_{k+1}}}{x_{\mathcal{S}_k}^{\mu_{k+1}}} = \frac{\exp(y_{\mathcal{S}_{k+1}})}{\exp(y_{\mathcal{S}_k})}, \quad (\text{B.23})$$

and hence

$$\frac{x_{\mathcal{S}_{k+1}}}{x_{\mathcal{S}_k}} = \frac{\exp(y_{\mathcal{S}_{k+1}}/\mu_{k+1})}{\exp(y_{\mathcal{S}_k}/\mu_{k+1})} = P_{\mathcal{S}_{k+1}|\mathcal{S}_k}(y) \quad (\text{B.24})$$

²¹Since $x_\alpha > 0$ for all $\alpha \in \mathcal{S}$, the multipliers for the corresponding inequality constraints all vanish by complementary slackness.

by the definition of the conditional logit choice model (55). Thus, by unrolling the chain

$$\begin{aligned} x_{\alpha|\mathcal{S}} &= \frac{x_{\alpha}}{x_{\mathcal{S}}} = \frac{x_{\mathcal{S}_L}}{x_{\mathcal{S}_{L-1}}} \cdot \frac{x_{\mathcal{S}_{L-1}}}{x_{\mathcal{S}_{L-2}}} \cdots \frac{x_{\mathcal{S}_{\ell+1}}}{x_{\mathcal{S}_{\ell}}} \\ &= P_{\mathcal{S}_L|\mathcal{S}_{L-1}}(y) \times P_{\mathcal{S}_{L-1}|\mathcal{S}_{L-2}}(y) \times \cdots \times P_{\mathcal{S}_{\ell+1}|\mathcal{S}_{\ell}}(y) \end{aligned} \quad (\text{B.25})$$

we obtain the nested expression

$$x_{\alpha} = x_{\mathcal{S}} \prod_{k=\ell}^{L-1} P_{\mathcal{S}_{k+1}|\mathcal{S}_k}(y) \quad \text{for all } \alpha \in \mathcal{S}. \quad (\text{B.26})$$

Hence, with $x_{\mathcal{S}} = 1$ (by the fact that $\text{supp}(x) = \mathcal{S}$), we finally conclude that

$$x_{\alpha} = \prod_{k=\ell}^{L-1} P_{\mathcal{S}_{k+1}|\mathcal{S}_k}(y) = P_{\alpha|\mathcal{S}}(y) \quad \text{for all } \alpha \in \mathcal{S}. \quad (\text{B.27})$$

Step 3: The maximal value of $(\text{Opt}_{\mathcal{S}})$. To obtain the value of the maximization problem $(\text{Opt}_{\mathcal{S}})$, we will proceed to substitute (B.27) in the expression (B.4) provided by Proposition B.1 for $h_{\mathcal{S}}(x)$. To that end, for all $k = \ell, \dots, L-1$ and all $\mathcal{S}_k \preceq_k \mathcal{S}$, the definition (B.3) of the conditional entropy gives:

$$\begin{aligned} h(x|\mathcal{S}_k) &= \mu_{k+1} x_{\mathcal{S}_k} \sum_{\mathcal{S}_{k+1} \triangleleft \mathcal{S}_k} x_{\mathcal{S}_{k+1}|\mathcal{S}_k} \log x_{\mathcal{S}_{k+1}|\mathcal{S}_k} && \# \text{ by definition} \\ &= \mu_{k+1} x_{\mathcal{S}_k} \sum_{\mathcal{S}_{k+1} \triangleleft \mathcal{S}_k} x_{\mathcal{S}_{k+1}|\mathcal{S}_k} \log \frac{\exp(y_{\mathcal{S}_{k+1}}/\mu_{k+1})}{\exp(y_{\mathcal{S}_k}/\mu_{k+1})} && \# \text{ by (B.24)} \\ &= x_{\mathcal{S}_k} \sum_{\mathcal{S}_{k+1} \triangleleft \mathcal{S}_k} x_{\mathcal{S}_{k+1}|\mathcal{S}_k} y_{\mathcal{S}_{k+1}} - x_{\mathcal{S}_k} y_{\mathcal{S}_k} && \# \text{ since } \sum_{\mathcal{S}_{k+1} \triangleleft \mathcal{S}_k} x_{\mathcal{S}_{k+1}|\mathcal{S}_k} = 1 \\ &= \sum_{\mathcal{S}_{k+1} \triangleleft \mathcal{S}_k} x_{\mathcal{S}_{k+1}} y_{\mathcal{S}_{k+1}} - x_{\mathcal{S}_k} y_{\mathcal{S}_k} \end{aligned} \quad (\text{B.28})$$

and hence

$$\begin{aligned} \sum_{\mathcal{S}_k \preceq_k \mathcal{S}} h(x|\mathcal{S}_k) &= \sum_{\mathcal{S}_k \preceq_k \mathcal{S}} \left[\sum_{\mathcal{S}_{k+1} \triangleleft \mathcal{S}_k} x_{\mathcal{S}_{k+1}} y_{\mathcal{S}_{k+1}} - x_{\mathcal{S}_k} y_{\mathcal{S}_k} \right] \\ &= \sum_{\mathcal{S}_{k+1} \preceq_{k+1} \mathcal{S}} x_{\mathcal{S}_{k+1}} y_{\mathcal{S}_{k+1}} - \sum_{\mathcal{S}_k \preceq_k \mathcal{S}} x_{\mathcal{S}_k} y_{\mathcal{S}_k}. \end{aligned} \quad (\text{B.29})$$

Thus, telescoping this last relation over $k = \ell, \dots, L$ and invoking Proposition B.1, we obtain:

$$\begin{aligned} h_{\mathcal{S}}(x) &= \sum_{\mathcal{S}' \preceq \mathcal{S}} h(x|\mathcal{S}') + \mu_{\mathcal{S}} x_{\mathcal{S}} \log x_{\mathcal{S}} && \# \text{ by Proposition B.1} \\ &= \sum_{k=\ell}^{L-1} \sum_{\mathcal{S}_k \preceq_k \mathcal{S}} h(x|\mathcal{S}_k) && \# \text{ collect parent classes} \\ &= \sum_{k=\ell}^{L-1} \left[\sum_{\mathcal{S}_{k+1} \preceq_{k+1} \mathcal{S}} x_{\mathcal{S}_{k+1}} y_{\mathcal{S}_{k+1}} - \sum_{\mathcal{S}_k \preceq_k \mathcal{S}} x_{\mathcal{S}_k} y_{\mathcal{S}_k} \right] && \# \text{ by (B.29)} \\ &= \langle y, x \rangle - x_{\mathcal{S}} y_{\mathcal{S}} \end{aligned} \quad (\text{B.30})$$

where, in the second line, we used the fact that the conditional entropy $h(x|\mathcal{S}_L)$ relative to any childless class $\mathcal{S}_L \in \mathcal{P}_L$ is zero by definition. Accordingly, substituting back to $(\text{Opt}_{\mathcal{S}})$

we conclude that

$$\text{val}(\text{Opt}_{\mathcal{S}}) = \langle y, x \rangle - h_{\mathcal{S}}(x) = x_{\mathcal{S}} y_{\mathcal{S}} = y_{\mathcal{S}}, \quad (\text{B.31})$$

as claimed.

Step 4: Differential representation of conditional probabilities. To prove the second part of the proposition, recall that the restricted entropy function $h_{|\mathcal{S}}$ is convex, and let

$$h_{|\mathcal{S}}^*(y) = \max_{x \in \Delta(\mathcal{A})} \{\langle y, x \rangle - h_{|\mathcal{S}}(x)\} \quad (\text{B.32})$$

denote its convex conjugate.²² By standard results in convex analysis (Rockafellar, 1970, Theorem 23.5), $h_{|\mathcal{S}}^*$ is differentiable in y and we have the Legendre identity:

$$x = \nabla h_{|\mathcal{S}}^*(y) \iff y \in \partial h_{|\mathcal{S}}(x) \iff x \in \arg \max_{x' \in \Delta(\mathcal{A})} \{\langle y, x' \rangle - h_{|\mathcal{S}}(x')\} \quad (\text{B.33})$$

Now, by (B.27), we have $x_{\alpha} = P_{\alpha|\mathcal{S}}(y)$ whenever x solves (Opt _{\mathcal{S}}) and hence, by Fermat's rule, whenever $y - \partial h_{|\mathcal{S}}(x) \ni 0$. Our claim then follows by noting that $h_{|\mathcal{S}}^*(y) = y_{\mathcal{S}}$ and combining the first and third legs of the equivalence (B.33). ■

The last remaining step is now trivial.

Proof of Theorem 3. Simply invoke Proposition B.2 with $\mathcal{S} \leftarrow \mathcal{A}$. ■

ACKNOWLEDGMENTS

This research was supported in part by the French National Research Agency (ANR) in the framework of the PEPR IA FOUNDRY project (ANR-23-PEIA-0003), the “Investissements d’avenir” program (ANR-15-IDEX-02), the LabEx PERSYVAL (ANR-11-LABX-0025-01), and MIAI@Grenoble Alpes (ANR-19-P3IA-0003). PM is also a member of the Archimedes/Athena RC, and was partially supported by project MIS 5154714 of the National Recovery and Resilience Plan Greece 2.0 funded by the European Union under the NextGenerationEU Program. WHS was supported by U.S. NSF Grant SES-1458992 and U.S. ARO Grant MSN201957.

REFERENCES

- Anderson, S. P., de Palma, A., and Thisse, J.-F. *Discrete Choice Theory of Product Differentiation*. MIT Press, Cambridge, MA, 1992.
- Auer, P., Cesa-Bianchi, N., Freund, Y., and Schapire, R. E. Gambling in a rigged casino: The adversarial multi-armed bandit problem. In *Proceedings of the 36th Annual Symposium on Foundations of Computer Science*, 1995.
- Auer, P., Cesa-Bianchi, N., and Fischer, P. Finite-time analysis of the multiarmed bandit problem. *Machine Learning*, 47:235–256, 2002.
- Beggs, A. W. On the convergence of reinforcement learning. *Journal of Economic Theory*, 122:1–36, 2005.
- Ben-Akiva, M. *Structure of Passenger Travel Demand Models*. PhD thesis, MIT, 1973.
- Ben-Akiva, M. and Lerman, S. R. *Discrete Choice Analysis: Theory and Application to Travel Demand*. MIT Press, Cambridge, 1985.
- Benaïm, M. and Weibull, J. W. Deterministic approximation of stochastic evolution in games. *Econometrica*, 71(3):873–903, May 2003.
- Binmore, K. and Samuelson, L. Muddling through: Noisy equilibrium selection. *Journal of Economic Theory*, 74(2):235–265, June 1997.
- Björnerstedt, J. and Weibull, J. W. Nash equilibrium and evolution by imitation. In Arrow, K. J., Colombarato, E., Perlman, M., and Schmidt, C. (eds.), *The Rational Foundations of Economic Behavior*, pp. 155–181. St. Martin's Press, New York, NY, 1996.

²²Note here that $h_{|\mathcal{S}}^*(y)$ is bounded from above by the convex conjugate $h_{\mathcal{S}}^*(y)$ of $h_{\mathcal{S}}(x)$ because the latter does not include the constraint $\text{supp}(x) \subseteq \mathcal{S}$.

- Börgers, T. and Mailath, G. J. Similarity-based learning and similarity equilibria. <https://www.youtube.com/watch?v=nvAqFT6WDTU>, 2020.
- Börgers, T. and Sarin, R. Learning through reinforcement and replicator dynamics. *Journal of Economic Theory*, 77(1):1–14, 1997.
- Bravo, M. and Mertikopoulos, P. On the robustness of learning in games with stochastically perturbed payoff observations. *Games and Economic Behavior*, 103(John Nash Memorial issue):41–66, May 2017.
- Coucheney, P., Gaujal, B., and Mertikopoulos, P. Penalty-regulated dynamics and robust learning procedures in games. *Mathematics of Operations Research*, 40(3):611–633, August 2015.
- Debreu, G. “Individual choice behavior: A theoretical analysis” by R. Duncan Luce. *American Economic Review*, 50(1):186–188, March 1960.
- Duvocelle, B., Mertikopoulos, P., Staudigl, M., and Vermeulen, D. Multi-agent online learning in time-varying games. *Mathematics of Operations Research*, 48(2):914–941, May 2023.
- Erev, I. and Roth, A. E. Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *American Economic Review*, 88:848–881, 1998.
- Farrell, J. and Klemperer, P. Switching costs and network effects. In Armstrong, M. and Porter, R. (eds.), *Handbook of Industrial Organization*, volume 3, pp. 1967–2072. Elsevier, Amsterdam, 2007.
- Friedman, D. Evolutionary games in economics. *Econometrica*, 59(3):637–666, 1991.
- Fudenberg, D. and Imhof, L. A. Imitation processes with small mutations. *Journal of Economic Theory*, 131:251–262, 2006.
- Fudenberg, D. and Imhof, L. A. Monotone imitation dynamics in large populations. *Journal of Economic Theory*, 140(1):229–245, May 2008.
- Fudenberg, D. and Kreps, D. M. Learning mixed equilibria. *Games and Economic Behavior*, 5(320–367), 1993.
- Fudenberg, D. and Levine, D. K. Consistency and cautious fictitious play. *Journal of Economic Dynamics and Control*, 19(5–7):1065–1089, 1995.
- Fudenberg, D. and Levine, D. K. Conditional universal consistency. *Games and Economic Behavior*, 29(1): 104–130, 1999.
- Giannou, A., Vlatakis-Gkaragkounis, E. V., and Mertikopoulos, P. Survival of the strictest: Stable and unstable equilibria under regularized learning with partial information. In *COLT ’21: Proceedings of the 34th Annual Conference on Learning Theory*, 2021.
- Gilboa, I. and Schmeidler, D. Case-based decision theory. *Quarterly Journal of Economics*, 110(3):605–639, August 1995.
- Hadikhanloo, S., Laraki, R., Mertikopoulos, P., and Sorin, S. Learning in nonatomic games, Part I: Finite action spaces and population games. *Journal of Dynamics and Games*, 9(4, William H. Sandholm memorial issue):433–460, October 2022.
- Helbing, D. A mathematical model for behavioral changes by pair interactions. In Haag, G., Mueller, U., and Troitzsch, K. G. (eds.), *Economic Evolution and Demographic Change: Formal Models in Social Sciences*, pp. 330–348. Springer, Berlin, 1992.
- Hirsch, M. W., Smale, S., and Devaney, R. L. *Differential Equations, Dynamical Systems, and an Introduction to Chaos*. Elsevier, London, UK, 2 edition, 2004.
- Hofbauer, J. and Sandholm, W. H. Stable games and their dynamics. *Journal of Economic Theory*, 144(4): 1665–1693, July 2009.
- Hofbauer, J. and Sigmund, K. *The Theory of Evolution and Dynamical Systems*. Cambridge University Press, 1988.
- Hofbauer, J., Schuster, P., and Sigmund, K. A note on evolutionarily stable strategies and game dynamics. *Journal of Theoretical Biology*, 81(3):609–612, 1979.
- Hofbauer, J., Sorin, S., and Viossat, Y. Time average replicator and best reply dynamics. *Mathematics of Operations Research*, 34(2):263–269, May 2009.
- Hopkins, E. Two competing models of how people learn in games. *Econometrica*, 70(6):2141–2166, November 2002.
- Jehiel, P. Analogy-based expectation equilibrium. *Journal of Economic Theory*, 123(2):81–104, August 2005.
- Jehiel, P. and Samet, D. Learning to play games in extensive form by valuation. *Journal of Economic Theory*, 124(2):129–148, October 2005.
- Jehiel, P. and Samet, D. Valuation equilibrium. *Theoretical Economics*, 2(2):163–185, 2007.

- Katz, M. L. and Shapiro, C. Systems competition and network effects. *The Journal of Economic Perspectives*, 8(2):93–115, 1994.
- Legacci, D., Mertikopoulos, P., and Pradelski, B. S. R. A geometric decomposition of finite games: Convergence vs. recurrence under exponential weights. In *ICML '24: Proceedings of the 41st International Conference on Machine Learning*, 2024.
- Leslie, D. S. and Collins, E. J. Individual Q -learning in normal form games. *SIAM Journal on Control and Optimization*, 44(2):495–514, 2005.
- Littlestone, N. and Warmuth, M. K. The weighted majority algorithm. *Information and Computation*, 108(2):212–261, 1994.
- Luce, R. D. Semiordeers and a theory of utility discrimination. *Econometrica*, 24(2):178–191, April 1956.
- Luce, R. D. *Individual Choice Behavior: A Theoretical Analysis*. Wiley, New York, 1959.
- Martin, M., Mertikopoulos, P., Rahier, T., and Zenati, H. Nested bandits. In *ICML '22: Proceedings of the 39th International Conference on Machine Learning*, 2022.
- McFadden, D. Modelling the choice of residential location. In Karlqvist, A., Lundqvist, L., Snickars, F., and Weibull, J. W. (eds.), *Spatial Interaction Theory and Planning Models*. North Holland, Amsterdam, 1978.
- McFadden, D. Economic choices. *American Economic Review*, 91:351–378, 2001.
- McFadden, D. L. The measurement of urban travel demand. *Journal of Public Economics*, 3(4):303–328, 1974a.
- McFadden, D. L. Conditional logit analysis of qualitative choice behavior. In Zarembka, P. (ed.), *Frontiers in Econometrics*, pp. 105–142. Academic Press, New York, NY, 1974b.
- McFadden, D. L. Econometric models of probabilistic choice. In Manski, C. F. and McFadden, D. L. (eds.), *Structural Analysis of Discrete Data with Econometric Applications*, pp. 198–272. MIT Press, Cambridge, MA, 1981.
- Mengel, F. Learning across games. *Games and Economic Behavior*, 74(2):601–619, March 2012.
- Mertikopoulos, P. and Moustakas, A. L. The emergence of rational behavior in the presence of stochastic perturbations. *The Annals of Applied Probability*, 20(4):1359–1388, July 2010.
- Mertikopoulos, P. and Sandholm, W. H. Learning in games via reinforcement and regularization. *Mathematics of Operations Research*, 41(4):1297–1324, November 2016.
- Mertikopoulos, P. and Sandholm, W. H. Riemannian game dynamics. *Journal of Economic Theory*, 177:315–364, September 2018.
- Mertikopoulos, P. and Viossat, Y. Survival of dominated strategies under imitation dynamics. *Journal of Dynamics and Games*, 9(4, William H. Sandholm memorial issue):499–528, October 2022.
- Mertikopoulos, P. and Zhou, Z. Learning in games with continuous action sets and unknown payoff functions. *Mathematical Programming*, 173(1-2):465–507, January 2019.
- Mertikopoulos, P., Papadimitriou, C. H., and Piliouras, G. Cycles in adversarial regularized learning. In *SODA '18: Proceedings of the 29th annual ACM-SIAM Symposium on Discrete Algorithms*, 2018.
- Mertikopoulos, P., Hsieh, Y.-P., and Cevher, V. A unified stochastic approximation framework for learning in games. *Mathematical Programming*, 203:559–609, January 2024.
- Nachbar, J. H. Evolutionary selection dynamics in games. *International Journal of Game Theory*, 19:59–89, 1990.
- Posch, M. Cycling in a stochastic learning algorithm for normal form games. *Journal of Evolutionary Economics*, 7:193–207, 1997.
- Rockafellar, R. T. *Convex Analysis*. Princeton University Press, Princeton, NJ, 1970.
- Rustichini, A. Optimal properties of stimulus-response learning models. *Games and Economic Behavior*, 29(1-2):244–273, 1999.
- Samuelson, L. and Zhang, J. Evolutionary stability in asymmetric games. *Journal of Economic Theory*, 57:363–391, 1992.
- Sandholm, W. H. Potential games with continuous player sets. *Journal of Economic Theory*, 97:81–108, 2001.
- Sandholm, W. H. *Population Games and Evolutionary Dynamics*. MIT Press, Cambridge, MA, 2010a.
- Sandholm, W. H. Pairwise comparison dynamics and evolutionary foundations for Nash equilibrium. *Games*, 1:3–17, 2010b.

- Sandholm, W. H. Population games and deterministic evolutionary dynamics. In Young, H. P. and Zamir, S. (eds.), *Handbook of Game Theory IV*, pp. 703–778. Elsevier, 2015.
- Schlag, K. H. Why imitate, and if so, how? A boundedly rational approach to multi-armed bandits. *Journal of Economic Theory*, 78(1):130–156, 1998.
- Shalev-Shwartz, S. Online learning and online convex optimization. *Foundations and Trends in Machine Learning*, 4(2):107–194, 2011.
- Shalev-Shwartz, S. and Singer, Y. Convex repeated games and Fenchel duality. In *NIPS' 06: Proceedings of the 19th Annual Conference on Neural Information Processing Systems*, pp. 1265–1272. MIT Press, 2006.
- Shub, M. *Global Stability of Dynamical Systems*. Springer-Verlag, Berlin, 1987.
- Sorin, S. and Wan, C. Finite composite games: Equilibria and dynamics. *Journal of Dynamics and Games*, 3(1):101–120, January 2016.
- Steiner, J. and Stewart, C. Contagion through learning. *Theoretical Economics*, 3:431–458, 2008.
- Taylor, P. D. and Jonker, L. B. Evolutionary stable strategies and game dynamics. *Mathematical Biosciences*, 40(1-2):145–156, 1978.
- Verboven, F. The nested logit model and representative consumer theory. *Economics Letters*, 50:57–63, 1996.
- Vovk, V. G. Aggregating strategies. In *COLT '90: Proceedings of the 3rd Workshop on Computational Learning Theory*, pp. 371–383, 1990.
- Weibull, J. W. *Evolutionary Game Theory*. MIT Press, Cambridge, MA, 1995.