A Representative-Sampling Model of Stochastic Choice*

Tuval Danenberg[†] and Ran Spiegler[‡]

April 17, 2025

Abstract

An agent facing a binary choice uses sampling to learn about payoffs. Each sample point carries Gaussian noise. The number of sample points about an alternative is proportional to its choice frequency. The agent chooses the best-performing alternative in the sample, ignoring sampling error. To account for sample-size endogeneity, we introduce an equilibrium concept for stochastic choice. The equilibrium effect favors the intrinsically inferior alternative, such that its choice frequency vanishes extremely slowly with total sample size. We also analyze how choices vary with the coarseness of the agent's sampling data, and illustrate how to extend this approach to games.

*This paper supersedes a previous, longer version titled "Representative Sampling Equilibrium". Financial support from UKRI Frontier Research Grant no. EP/Y033361/1 and from the Sapir Center is gratefully acknowledged. We thank Jose Apesteguia, Ian Ball, Itamar Bellaiche, Alex Clyde, Drew Fudenberg, Nathan Hancart, Paola Manzini, Meg Meyer, Yuval Salant, Tomasz Strzalecki, seminar participants at NHH and MIT, and the editor and referees of this journal, for helpful feedback. This paper was edited by Jay Lu.

[†]MIT. E-mail: tuvaldan@mit.edu.

[‡]Tel Aviv University and University College London. https://www.ranspiegler.sites.tau.ac.il/. E-mail: rani@tauex.tau.ac.il.

1 Introduction

Additive Random Utility (ARU) is probably the most familiar modeling approach to stochastic choice (see Strzalecki (2023) for a pedagogical exposition). According to the ARU model, each choice alternative a carries an intrinsic utility $u(a)$. When the decision-maker (DM) faces a choice between alternatives, she evaluates a by $u(a)+\varepsilon$, where ε represents independently distributed noise. This noise term is commonly interpreted as non-systematic population-wide variation in the motivations of DMs, or within a single DM across choice situations. In both cases, ε represents uncertainty of an *outside observer*.

Another interpretation of ARU is that $u(a)+\varepsilon$ represents a noisy signal obtained by the DM *herself*, lacking direct access to her intrinsic valuation of each alternative. This process may involve introspection — for instance, trying to retrieve past experiences from memory. Alternatively, it may involve physical sampling of other agents' experiences (asking friends, reading product reviews). The DM *naively extrapolates* from her noisy signal: she regards the signal as a perfect predictor of the intrinsic value of a and chooses the alternative that maximizes $u(a)+\varepsilon$ in her sample.

This interpretation of random choice harks back to Thurstone's (1927) model of noisy perception, according to which perceived stimulus is the sum of true stimulus and normally distributed noise. In one of the examples that motivated Thurstone's analysis, an agent asked to identify the heavier of two objects generates Gaussian weight signals and picks the object with the higher signal. The naive-sampling interpretation of ARU extends this idea from perception of external objects to perception of subjective preferences. We refer to this interpretation as naive sampling because it describes the DM as a "*naive frequentist*" who obtains noisy additive signals of choice alternatives' intrinsic value and takes these signals at face value, neglecting sampling error.

However, this description raises a natural question: Shouldn't alternatives that are chosen more frequently generate more precise signals? Suppose the error term ε captures the noisiness of an introspective process by which the DM tries to access the intrinsic value of a choice alternative. Then, when the DM chooses an alternative more frequently, she is likely to have an easier time retrieving memories of that alternative. Or consider the physical-sampling interpretation. When an alternative is chosen more frequently in the relevant population, the DM can draw on a bigger sample of peers' experiences with this alternative.

Under both interpretations, the variance of the error term should *decrease* with its popularity. This dependence creates a feedback effect: Choice frequencies depend on DMs' subjective evaluations of alternatives, and yet these very evaluations are sensitive to choice frequencies. This feedback effect suggests a need for an *equilibrium* concept of single-agent stochastic choice. To capture this idea, we modify the standard ARU model. Conventionally, we assume that the DM observes the value of each choice alternative with additive Gaussian noise. We depart from the standard model by assuming that the *variance* of this noise depends on *the frequency with which x is chosen*.

Specifically, we consider a setting in which the DM chooses between two alternatives, A and B. The DM obtains a sample of size n, consisting of $nq(A)$ and $nq(B)$ observations about A and B, where $q(z)$ is the choice frequency of alternative z. Thus, the DM's sample is *representative*. Each sample point about alternative z generates an observed utility of $u(z) + \varepsilon$, where ε is an independent draw from a *normal distribution* with mean zero and variance σ^2. Thus, the average utility in the DM's sample for alternative z has mean $u(z)$ and variance $\sigma^2/nq(z)$. In keeping with the "naive frequentism" idea, the DM chooses the alternative with the highest average utility in her sample. In a *representative sampling equilibrium* (RSE), the choice frequencies that result from this procedure match q.

Under the physical-sampling interpretation, representative sampling can be taken literally, modeling a form of experimentation in which the DM *deliberately* ensures that the composition of her sample matches the relevant population, somewhat in the manner of political pollsters. However, we prefer to think of representative sampling as a "mean field" approximation of passive observational learning, where the DM faces a random sample drawn from the equilibrium distribution. Under the introspective interpretation, representative sampling captures an internal process of evaluation. As the DM becomes more familiar with an alternative, her introspective process generates a more precise signal. For both interpretations, the representative-sampling approximation makes the model tractable while preserving the feature that frequently chosen alternatives generate more precise signals.

For a concrete example of the introspective interpretation, consider an agent choosing between two brands of beer, A and B. Suppose the agent would always derive greater expected pleasure from beer B. However, she does not know her taste for beer well enough to recognize this. Instead, she relies on her personal memory of previous experiences drinking these beers. These experiences are noisy, for example due to variation in dish pairings or atmosphere. Importantly, the composition of the sample will reflect the agent's previous choices: if she mostly drank B in the past, she will have a more precise understanding of her pleasure from this beer. The agent's memory is bounded. As time goes by, she accumulates new experiences and forgets others. In a steady state, the probability that the agent opts for B should equal the historical frequency of choosing it. RSE captures this notion of a steady state.

As to the physical-sampling interpretation, for a concrete example think of an agent choosing between two hotels. Prior to making her choice, the agent reads online reviews or asks friends who visited one of the hotels about their experiences. Suppose that the description of these experiences is complete, as if they happened to the agent herself (such that we can abstract from the usual inferential challenges of social learning). The noise might be due to objective variation in service quality at the hotels. The sample size for each hotel will depend on its popularity, such that the agent will have a more precise impression of more popular hotels.

The key insight of our model is that naive inference from representative samples introduces an *equilibrium force that favors inferior alternatives*. In the beer example, the assumption that the one brand of beer is intrinsically inferior (according to the DM's true underlying taste) leads the DM to have fewer sample points about this beer, which makes her assessment of it noisier. The DM's naive-frequent inference implies that a noisy assessment favors an intrinsically inferior alternative. This generates an equilibrium effect that magnifies the inferior alternative's choice frequency.

After establishing existence, uniqueness, and monotonicity results for RSE, our main result addresses the implications of the basic insight for how choice frequencies depend on the sample size n. Not only does representative sampling increase the equilibrium frequency of the inferior alternative relative to the rational or uniform-sample benchmarks, but the *rate* at which this frequency vanishes with n is *extremely slow*. With uniform sampling, the probability of choosing the superior alternative is approximately $\Phi(\sqrt{n})$, where Φ is the *cdf* of the standard normal distribution. With RSE, for any positive k there exists n after which this probability falls below $\Phi(n^k)$.

Later in the paper, we consider extensions of the basic model. First, we assume the DM has multiple types, which differ in the magnitude of the utility difference $u(B) - u(a)$ while agreeing on its sign. Types are partitioned into "intervals", such that each type's sample is restricted to types in the same interval. We extend the existence, uniqueness and monotonicity results of the basic model, and carry out comparative statics with respect to the partition's *coarseness*. When intrinsic taste differences are not too large, *finer* partitions lead to *larger* RSE choice errors. In other words, more detailed data can lower decision quality.

Next, we extend RSE to games, and show that the favoring-inferior-alternatives effect of RSE has non-trivial implications for cooperation patterns in static and dynamic versions of the Prisoner's Dilemma. Finally, we consider a variant on RSE in which the DM's inference from the sample is Bayesian under a symmetric prior, taking into account the representation of each alternative in the sample. We explain why even this Bayesian version is not a "proper" rational-choice benchmark. We illustrate numerically that this (analytically intractable) variant generates multiple asymmetric equilibria even when the two alternatives are intrinsically identical.

2 Related Literature

Despite the intuitive appeal of the naive-sampling interpretation of ARU and its historical connection to Thurstone (1927), it has not received much attention in the literature on stochastic individual choice. Yet, the naive-sampling approach to random choice has been developed in other contexts. Osborne and Rubinstein (1998) introduced the game-theoretic concept of $S(K)$ equilibrium, in which each player samples each available strategy K (independent) times and chooses the best-performing strategy in her sample. Osborne and Rubinstein (2003) studied a variant on this concept (in the context of a voting model), in which each player best-replies to a finite sample drawn from her opponents' strategies. Spiegler (2006a,b) examined price competition models in which consumers evaluate products using the $S(K)$ procedure. Sethi (2000) formalized Osborne and Rubinstein's dynamic interpretation of $S(1)$ equilibrium.

Osborne and Rubinstein (1998, 2003) assumed that players regard their sample as a noiseless estimate of the distribution from which it is drawn. This is what we referred to as "naive frequentist" inference, which this paper assumes as well. Salant and Cherry (2020) extended the sampling-based equilibrium approach to a more general class of statistical inference procedures, and introduced new methods for analyzing equilibria. Unlike the present paper, Salant and Cherry (2020) maintained Osborne and Rubinstein's assumption that sample size is an exogenous parameter.

Naive frequentist inference from random samples (which involves neglect of sampling error) is related to what Tversky and Kahneman (1971) called "the law of small numbers" — namely, treating small samples as if they are perfectly representative of the distribution they are drawn from. The idea that people take sample averages at face value and inadequately incorporate sample size has received corroboration both in experimental settings (e.g., Orbrecht et al. (2007)) and in studies of users' responses to online reviews (e.g., de Langhe et al. (2016)). The observation that frequent use of products leads to more precise information about their quality has been considered from a very different perspective in the theoretical IO literature — e.g., see Shapiro (1983) on the pricing of experience goods.

The literature also contains models of random choice arising from Bayesian learning about the values of alternatives — see Strzalecki (2023) for a textbook treatment. Lu (2016) axiomatizes such a model, and demonstrates how agents' private information can be identified from random choice data. Natenzon (2019) uses a related framework to explain choice anomalies such as IIA violation. Fudenberg et al. (2024) show that a model of agents learning from finite memory generates behavior that falls within the random choice representation in Lu (2016). In a game-theoretic context, Goncalves (2020) presented a sequential-sampling-based solution concept that adheres to Bayesian rationality, and used this concept to produce joint predictions about choice behavior and decision times.

The physical-sampling interpretation of our model links it to the literatures on word-of-mouth learning (e.g., Ellison and Fudenberg (1995), Banerjee and Fudenberg (2004)) and learning in social networks (e.g., Golub and Jackson (2012)). Unlike this paper, both literatures involve explicitly dynamic models. Like us, Banerjee and Fudenberg (2004) assume that the process of social learning involves representative samples. However, they assume that agents draw Bayesian inferences from noisy observations of their predecessors' payoffs (as well as their observed choices). An important distinction between our paper and these works (and social-learning models in the tradition of Bikchandani et al. (1992) and Banerjee (1992)), is that agents in our model do not draw inferences from observed choices as such.

3 A Model

A DM faces a choice between two alternatives, denoted A and B. The DM's expected utility from choosing an alternative $z \in \{A, B\}$ is $u(z)$. We assume that alternative B is superior, so $u(A) < u(B)$. Let $q(z)$ be the probability that the DM chooses z. We will often use the abbreviated notation $q = q(B)$. In our model, the DM does not know the expected utilities, and q is a consequence of her attempt to learn about them through sampling. The DM's total sample size is a positive integer n. Her estimate of $u(z)$ is independently and normally distributed as follows:

$$\hat{u}(z) \sim N\left(u(z), \frac{\sigma^2}{nq(z)}\right) \tag{1}$$

where $\sigma^2 > 0$ is the variance of a sample point from any alternative.

The motivation behind (1) is that the DM has a representative sample, consisting of $nq(z)$ observations of alternative z. Each observation carries Gaussian noise with variance σ^2. Thus, to recall our hotel example from the Introduction, if the fraction of consumers in the general population who choose hotel B is 0.6 and our DM has a sample of total size $n = 10$, then she will have six observations about alternative B and four observations about alternative A. Of course, in general $nq(z)$ need not be an integer: As explained in the Introduction, we conceive of (1) as an approximation of a value estimate that arises from a random sample drawn from the general consumer population.

When $q(z) = 0$, $\hat{u}(z)$ is ill-defined because it involves infinite variance. To handle this, we treat $N(0, \infty)$ as a distribution satisfying $\Pr(x \leq c) = \frac{1}{2}$ for every c.

Definition 1. *A probability q of choosing alternative B is a **representative-sampling equilibrium (RSE)** if*

$$q = \Pr(\hat{u}(B) - \hat{u}(A) > 0)$$

where this probability is calculated according to (1).

The idea behind RSE is as follows. Before choosing an action, the DM samples the utility of each alternative. An alternative's representation in her sample matches its choice frequency. The DM is a "naive frequentist" who takes sample outcomes at face value. That is, she regards the sample average $\hat{u}(z)$ as an accurate representation of her underlying expected utility from choosing z, ignoring sampling error.

By the assumption that $\hat{u}(A)$ and $\hat{u}(B)$ are distributed according to (1), their difference is distributed as follows:

$$\hat{u}(B) - \hat{u}(A) \sim N\left(u(B) - u(A), \frac{\sigma^2}{nq(1-q)}\right) \tag{2}$$

Denote $t = u(B) - u(A)$. Our assumption that B is superior implies $t > 0$. It is clear from (2) that the value of σ can be *normalized* without loss of generality, because we can rescale t. From now on, we set $\sigma = 1$. Consequently, the equilibrium condition can be rewritten as

$$q = Pr\left[N\left(0, \frac{1}{nq(1-q)}\right) < t\right]$$

or, equivalently,

$$q = \Phi\left(t\sqrt{nq(1-q)}\right) \tag{3}$$

where Φ is the *cdf* of the standard normal distribution (we invoke this conventional notation consistently throughout the paper).

We will use (3) as our *working definition* of RSE in Section 4. This definition immediately implies that in any RSE, the DM chooses the superior alternative (i.e., the z with the higher $u(z)$) with probability greater than $\frac{1}{2}$. It is not surprising that due to sampling errors, the inferior alternative is also chosen with positive probability in RSE.

Yet, how big is the DM's choice error? A central theme of this paper is that representative samples (coupled with naive-frequentist inference) magnify the probability of errors. Specifically, when the choice distribution is skewed (i.e., when q is close to zero or one), the variance of $\hat{u}(B) - \hat{u}(A)$ is large, and this introduces an equilibrium counter-force toward a less skewed distribution, namely larger choice errors.

The intuition that noisy measurement favors inferior alternatives is quite basic and does not rely on the Gaussian noise structure of our model. For example, suppose that the DM observes the value of each alternative z with independent exogenous noise that takes the two values e and $-e$ with equal probability, where $e > 0$. Then, if $e < |u(B) - u(A)|/2$, the superior alternative will always be chosen, whereas when $e > |u(B) - u(A)|/2$, the inferior alternative will be chosen with probability $\frac{1}{4}$. This is a simple illustration that inferior alternatives can benefit from noisy assessment when DMs naively follow their signal.[1] Of course, the simple binary noise structure in this example cannot be obtained as the average of multiple independent and identically distributed observations. In fact, a key reason for using a Gaussian noise structure is that it is well-suited for distributions of sample averages, and in particular distributions of differences between averages of samples of variable size. Thus, while the Gaussian structure is not necessary for the effect that noise favors inferior alternatives, it enables thorough exploration of this effect as generated by the representative-sampling assumption.

Comment: What does the DM observe?
One interpretation of $\hat{u}(z)$ is that it represents naive-frequentist inference from a physical sample of *other people*'s choices. One way to make this interpretation consistent is to assume that the DM has enough detail about each data point to learn what her utility would be if this were *her own* experience. Using our hotel example from the Introduction, each sample point consists of a complete description of a friend's consumption experience. Since the experience contains random elements (room allocation, staffing), it is a noisy signal of the DM's own expected utility if she chooses the same hotel. However, since the DM has access to the friend's full experience, she knows what her utility from that same experience would be.

4 Analysis

We begin our analysis with an elementary result.

Proposition 1. *There is a unique RSE.*

This is a special case of a more general result we present in Section 5, whose proof (like all proofs in this paper) appears in the Appendix.

A basic observation is that the probability of choosing the superior alternative B increases in the utility difference in favor of B and in the sample size.

Remark 1. *The RSE value of q is strictly increasing in t and n.*

[1] For illustrations of the implications of this effect for market competition, see Spiegler (2011, Ch. 7) and Szech (2011).

For our next exercise, let $q(n)$ denote the RSE value of q when the sample size is n. By Remark 1, $q(n)$ is strictly increasing. The next result shows that choice errors vanish as n tends to infinity.

Proposition 2. $\lim_{n\to\infty} q(n) = 1$.

Before we address the speed of this convergence, let us consider the benchmark case of a *uniform sample*, where each alternative is sampled $n/2$ times. This variant shares the sampling-based account of random choice, while suppressing the idea that choice frequencies affect signal precision. In the uniform-sample case, the probability of choosing B is given by

$$r(n) = \Phi\left(\frac{t}{2}n^{\frac{1}{2}}\right) \tag{4}$$

This can be viewed as a normal approximation of Osborne and Rubinstein's (1998) $S(K)$ procedure, where $K = n/2$.

Formula (4) has two noteworthy features. First, it lacks the equilibrium effect that arises from representative sampling. Second, since $\sqrt{q(1-q)} \le \frac{1}{2}$ for any $q \in (0,1)$, $r(n)$ assigns higher probability to the favored alternative than $q(n)$.

Of course, $r(n)$ increases with n and converges to one as $n \to \infty$. However, $r(n)$ differs from $q(n)$ in the *speed* of convergence. Our next result demonstrates that $q(n)$ increases *much more slowly* than $r(n)$. For convenience and without loss of generality, we fix $t = 1$.

Proposition 3. *Let $t = 1$. For every $k > 0$, there exists $n(k)$ such that for every integer $n \ge n(k)$: $q(n) \le \Phi(\frac{1}{2}n^k)$*

In the uniform-sample case, $r(n)$ increases with n like $\Phi(\sqrt{n})$. By comparison, in the representative sample case, $q(n)$ increases with n more slowly than $\Phi(n^k)$ for *any* k, however small (and in particular, smaller than $\frac{1}{2}$). Thus, the equilibrium forces introduced by representative sampling have a qualitative effect on the DM's choice behavior, even when n is large.

Figure 1 illustrates this comparison for $t = 1$. Figure 1(a) focuses on the range $n < 100$, while Figure 1(b) zooms out to $n < 500$ (and also plots $\Phi(\frac{1}{2}n^{1/4})$). As we can see, the uniform-case specification exhibits fast convergence — e.g., $r(30) \approx 0.997$.

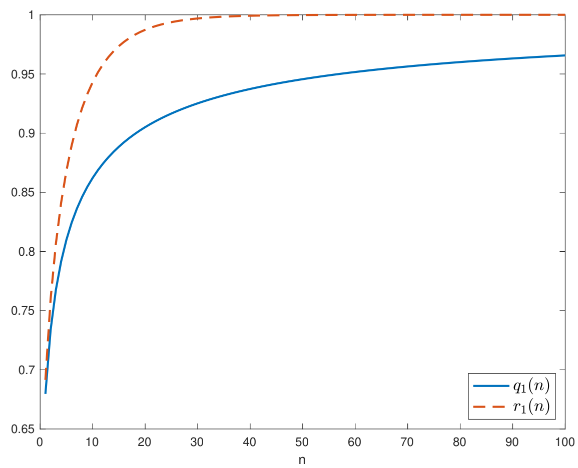

Figure 1(a)

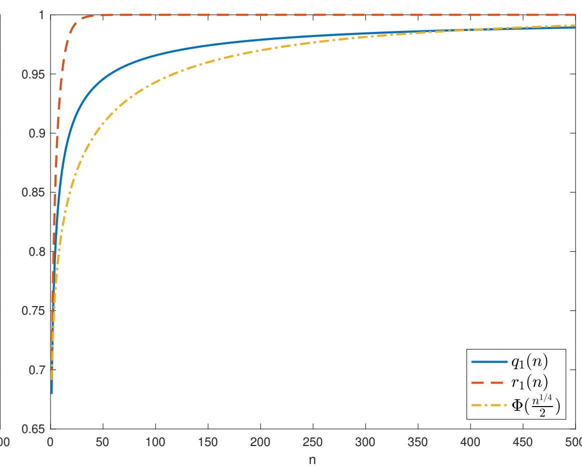

Figure 1(b)

In contrast, the RSE prediction is $q(30) \approx 0.925$. Considering that $t = 1$ represents a utility difference of one standard deviation between the alternatives (recall that $\sigma = 1$), this is a significant choice error. Moreover, convergence is very slow such that from around $n = 400$, $q(n) < \Phi(\frac{1}{2}n^{1/4})$. In this sense, RSE predicts larger choice errors than the uniform-sampling model.

Since t and $\sqrt{n}$ appear multiplicatively on the R.H.S. of (3), Proposition 3 also means that RSE choice probabilities vary slowly with t. By comparison, in a standard ARU with Gaussian noise (namely, a Probit model), the probability of choosing the superior alternative would display fast convergence to 1 as we raise t, because of the normal distribution's thin tail. Thus, even though the underlying noise in our model is Gaussian, choice probabilities exhibit fat-tail behavior, as a result of the equilibrium effects of representative sampling. This is the key behavioral distinction between RSE and the Probit model.

5 Getting Data from "Similar" Types

In many of the real-life situations that motivate our physical sampling interpretation, people do not get their data from a representative sample of agents who are *identical* to them, but rather from a population of *similar* agents. Even when the sample consists entirely of the DM's own experiences, these may involve *choice situations* that are similar but not identical to the one she is facing. To capture this, we extend the model by assuming that the intrinsic utility difference between the alternatives is not constant.

Let $T \subset \mathbb{R}_{++}$ be a finite set of DM types. Let $\mu \in \Delta(T)$ represent a distribution over types, where μ_t denotes the fraction of DMs of type t. The DM's expected utility from choosing an alternative $z \in \{A, B\}$ given her type $t \in T$ is denoted $u(z,t)$. Henceforth, we identify t with $u(B,t) - u(A,t)$, namely the intrinsic utility difference between the two alternatives. Let $q_t(z)$ be the probability that a DM of type t chooses z. We will occasionally use the abbreviated notation $q_t = q_t(B)$.

Let Π be a partition of T, where $\Pi(t)$ denotes the partition cell that includes t. We assume that Π is an *interval partition*: If $\Pi(t) = \Pi(t')$ and $t < t'' < t'$, then $\Pi(t'') = \Pi(t)$. We define the following binary relation over Π: $\pi' \succ \pi$ if and only if $t' > t$ for every $t \in \pi$, $t' \in \pi'$. Since Π is an interval partition, $\succ$ is a linear order. The average frequency of choosing z among types in $\Pi(t)$ is

$$\bar{q}_{\Pi(t)}(z) = \frac{\sum_{t \in \Pi(t)} \mu_t q_t(z)}{\sum_{t \in \Pi(t)} \mu_t} \tag{5}$$

We will occasionally use the abbreviated notation $\bar{q}_{\Pi(t)} = \bar{q}_{\Pi(t)}(B)$.

One interpretation of Π is that it captures *coarse sample data*, in the spirit of Jehiel's (2005) notion of analogy partitions. Under the physical-sampling take on our model, the DM tends to learn the outcome of choices by other agents who are *like her*, in the sense that they share certain characteristics with her. Under the introspective take, the DM accesses similar situations she experienced in the past.

The extension of RSE to this setting is straightforward. As before, the DM's total sample size is a positive integer n. The DM's estimate of $u(z,t)$ is independently and normally distributed as follows:

$$\hat{u}(z,t) \sim N\left(u(z,t), \frac{\sigma^2}{n\bar{q}_{\Pi(t)}(z)}\right) \tag{6}$$

This reflects the assumption that the composition of the DM's sample is determined by the *average* choice frequencies of types in her partition cell.

Definition 2. *A profile $q = (q_t)_{t \in T}$ is an RSE if for every $t \in T$,*

$$q_t = \Pr(\hat{u}(B,t) - \hat{u}(A,t) > 0)$$

where this probability is calculated according to (6).

Setting $\sigma = 1$ and following the same derivation as in Section 3, the condition for RSE can be stated equivalently as follows. For every t,

$$q_t = \Phi\left(t\sqrt{n\bar{q}_{\Pi(t)}(1 - \bar{q}_{\Pi(t)})}\right) \tag{7}$$

Comment: The assumption that $T \subset R_{++}$

Since we identify t with the intrinsic utility difference between alternatives B and A, the assumption that all t's are positive means that all agents derive higher expected utility from alternative B and differ only in the magnitude of the utility difference. One interpretation for this restriction is that B is of *objectively higher quality* and DM types differ in how sensitive they are to quality differences. This assumption is used to prove the following uniqueness result. Finding conditions for equilibrium uniqueness when the sign of t is not constant is an open problem.

Proposition 1*. *There is a unique RSE.*

This result implies Proposition 1 as a special case. And, as in the basic model, $q_t > \frac{1}{2}$ for every $t \in T$ in RSE.

We now explore monotonicity properties of RSE in this extended case. First, in an RSE profile q, if $t' > t$ and $\Pi(t') = \Pi(t)$, then $q_{t'} > q_t$. To see why, note that if $\Pi(t') = \Pi(t)$, then $\bar{q}_{\Pi(t')} = \bar{q}_{\Pi(t)}$, such that the R.H.S of (7) is higher for t' than for t. However, monotonicity may fail for types that belong to different partition cells. In particular, it is possible that $t' > t$ and yet $q_{t'} < q_t$ in the unique RSE. To see why, note that in RSE, two opposing forces shape choice probabilities. On one hand, a higher type, which represents a greater underlying taste for B, is a force that increases the probability of choosing this alternative. On the other hand, suppose that $\Pi(t') \succ \Pi(t)$ and t' is at the lower end of its cell while t is at the upper end of its cell. Then, t' shares its cell with higher types that imply a high $\bar{q}_{\Pi(t')}$, whereas t shares its cell with lower types that imply a low $\bar{q}_{\Pi(t)}$. As a result, the sample size for alternative A will be smaller for type t', which implies a noisy estimate of the utility difference between the two alternatives. This force favors the inferior alternative A, and therefore lowers the probability of choosing B for t', relative to t. The net effect of these two forces is ambiguous.

The next result establishes monotonicity of $\bar{q}_\pi$, and implies Remark 1 as a special case.

Proposition 4. *If q is an RSE profile, then $\pi \succ \pi'$ implies $\bar{q}_\pi > \bar{q}_{\pi'}$.*

We now turn to the question of how the coarseness of Π affects the DM's behavior. First, we analyze the effect of splitting a partition cell into multiple sub-cells on the average behavior of types in the various sub-cells.

Proposition 5. *Consider two partitions Π and Π', such that Π' refines some cell T^* into a collection of sub-cells $\{T^1, ..., T^m\}$. Let q and q' be the RSE under Π and Π'. Then:*

(i) If $\bar{q}_{T^k} > \bar{q}_{T^*}$, then $\bar{q}'_{T^k} < \bar{q}_{T^k}$.

(ii) If $\bar{q}_{T^k} < \bar{q}_{T^*}$, then $\bar{q}'_{T^k} > \bar{q}_{T^k}$.

To understand this result, suppose that the original partition is fully coarse, and its refinement divides it into two groups. Suppose further that under the original coarse partition, the average propensity to consume the superior alternative in group 1 is above the overall average (such that group 2 is below the average). The result says that after the refinement, the average probability of consuming the superior alternative decreases in group 1 and increases in group 2. The intuition behind this result is that when members of group 1 stop learning from the choices of members of group 2, they have fewer sample points about the inferior product, which leads to a noisier assessment and therefore a lower probability of choosing the superior product.

We now show that as long as the types in T are not too far away from zero, a finer partition leads to a higher overall probability of taking the inferior action A. Denote

$$\bar{q}(\Pi) = \sum_{t \in T} \mu_t q_t(\Pi)$$

where $q_t(\Pi)$ is the RSE probability that type t chooses B under the partition Π.

Proposition 6. *Suppose $t\sqrt{n} \in (0,2)$ for every $t \in T$. Consider two interval partitions Π and Π', such that Π' is a refinement of Π. Then, $\bar{q}(\Pi') < \bar{q}(\Pi)$.*

This result establishes that when the intrinsic advantage of alternative B is not too large, a finer partition leads to a higher probability of choice mistakes. Under the "coarse data" interpretation, the result means that finer data has an adverse effect on average choice quality. The question of how the coarseness of Π affects average behavior for larger values of t remains open.

It can also be shown that under the same conditions, a finer partition has an adverse effect on average *welfare*. Intuitively, this is because Proposition 5 implies that refining the partition leads to a decrease (an increase) in the probability of choosing B among high (low) types. Proposition 6 shows that the decrease among the high types is greater than the increase among the low types. Since the welfare effects of a change in choice probability are larger for high types (whose bias in favor of B is stronger), Proposition 6 also implies an overall decrease in average welfare following the refinement.

6 Other Extensions

The previous section extended the basic model of Section 3 by considering multiple DM types with coarse samples. We now return to the single-type choice setting and briefly discuss two alternative extensions. First, we consider the implications of replacing the naive-frequentist DM with one who employs Bayesian inference. Second, we show how naive-frequentist RSE can be extended to non-binary choice problems.

6.1 RSE with Bayesian Inference

The definition of RSE in Section 3 was based on two ideas. First, the strength of the DM's signal regarding an alternative (measured by the inverse of the signal's variance) is proportional to the (endogenous) choice frequency of this alternative. Second, the DM's inference from her signal is "naive frequentist" — i.e., she takes the signal at face value and uses it as a prediction of the alternative's true value. In this subsection, we retain the first element but replace naive frequentism with Bayesian inference, which takes into account samples' size-dependent informational content. In this version of RSE, the DM realizes that smaller samples are less informative.

Recall that for $z \in \{A, B\}$, $u(z)$ is the expected utility from alternative z, and $q(z)$ is the probability that the DM chooses z. To define Bayesian inference, we need to specify a prior belief. Assume the DM's belief over each of the alternatives is given by an independent standard normal distribution, $N(0, 1)$. A single observation of alternative z is distributed according to $N(u(z), \sigma^2)$. The sample average of alternative z is $x(z) \sim N(u(z), \sigma^2(z))$, where $\sigma^2(z) = \sigma^2/nq(z)$. Given the sample average realization $x(z)$, the DM's posterior mean belief for alternative z is $m(z) = x(z)/(1+\sigma^2(z))$. Therefore,

$$m(z) \sim N\left(\frac{u(z)}{1+\sigma^2(z)}, \frac{\sigma^2(z)}{(1+\sigma^2(z))^2}\right)$$

such that

$$m(B) - m(A) \sim N\left(\frac{u(B)}{1+\sigma^2(B)} - \frac{u(A)}{1+\sigma^2(A)}, \frac{\sigma^2(B)}{(1+\sigma^2(B))^2} + \frac{\sigma^2(A)}{(1+\sigma^2(A))^2}\right)$$

A value $q = q(B)$ is a Bayesian RSE if $q = \Pr(m(B) - m(A) > 0)$, where this probability is computed for samples with a share q of alternative B. Equivalently,

$$q = \Phi\left(\frac{u(B)(1+\sigma^2(A)) - u(A)(1+\sigma^2(B))}{\sqrt{\sigma^2(B)(1+\sigma^2(A))^2 + \sigma^2(A)(1+\sigma^2(B))^2}}\right) \tag{8}$$

Comment: Is the Bayesian-RSE DM rational?
Even though this DM is Bayesian, her inference is inconsistent with rational expectations, once we think about the dynamic learning process that implicitly underlies the equilibrium concept. The Bayesian-RSE DM takes into account sample sizes when performing her inference from the sample averages. However, she does not take into account the reasoning that led to these sample sizes in the first place. In a dynamic learning model, a Bayesian DM with access to the outcomes of choices by n other agents would draw inferences not only from these outcomes but also from the choice frequencies in the sample. These frequencies reflect the other agents' own inferences from the samples at their disposal, which our DM does not observe. This is a standard feature of Bayesian social learning processes, which Bayesian RSE does not incorporate. Formulating a Bayesian-rational learning model that incorporates these inferences would require abandoning the static stochastic-choice framework, and specifying an explicit extensive-form game that describes the social-learning process.

Formula (8) is significantly more complex than formula (3), which defines our original notion of RSE (based on naive frequentism). First, while (3) only makes use of the difference $t = u(B) - u(A)$ between the two alternatives' intrinsic utility, the Bayesian version is a function of $u(A)$ and $u(B)$ individually (because their relation to the prior belief is also relevant). Second, the dependence of the R.H.S of (8) on q is significantly more complicated than in (3). This added complexity makes the model considerably less amenable to analytic characterization. In addition, numerical analysis reveals that unlike our basic RSE, the Bayesian version admits multiple equilibria. This feature appears even when $u(A) = u(B) = u$ — i.e., when the two alternatives are intrinsically identical. The basic definition of RSE admits a unique equilibrium $q = \frac{1}{2}$ in this case. While this continues to be an equilibrium under Bayesian RSE, asymmetric equilibria can emerge (note that whenever q is an equilibrium, $1 - q$ is an equilibrium as well). Some examples are:

u	$\frac{\sigma^2}{n}$	q
0.5	0.01	0.108211
0.9	0.23	0.355468
1.5	0.71	0.294828
2	1.05	0.10644

The intuition behind asymmetric equilibria is as follows. Suppose that average utilities in the DM's sample are $x(A) > x(B) > 0$. This leads to a stronger tendency to choose alternative A. As a result, the sample size about alternative B is smaller. A Bayesian DM, unlike the naive-frequentist one, reacts to this paucity of evidence by heavily discounting $x(B)$ toward the prior zero. This exacerbates the perceived difference between the two alternatives, which perpetuates the tendency to choose alternative A. Thus, asymmetric equilibria are possible in Bayesian RSE, even when the two alternatives are intrinsically identical. While potentially interesting by itself, this multiplicity of equilibria also means that the Bayesian version of RSE is less convenient for applications than our original version, in addition to the complications that arise from its highly intractable formula.

6.2 Non-Binary Choice

Extending the model of Section 3 to choice settings with more than two alternatives is conceptually straightforward. Consider a naive-frequentist DM choosing between alternatives $A_1, \ldots, A_n$. Let $u(A_i)$ denote her expected utility from A_i, and let her estimate of the utility from each alternative A_i be independently and normally distributed as follows:

$$\hat{u}(A_i) \sim N\left(u(A_i), \frac{\sigma^2}{nq(A_i)}\right)$$

where $q(A_i)$ is the probability that the DM chooses A_i, and σ^2 is the variance of a sample point from any alternative.

The profile $(q(A_i))_{i=1,\ldots,n}$ is an RSE if for all $i = 1, \ldots, n$,

$$q(A_i) = \Pr(\hat{u}(A_i) > \max_{j \neq i} \hat{u}(A_j))$$

Fixing $\sigma^2 = 1$, the RSE equations can be written as

$$q(A_i) = \int_{-\infty}^{+\infty} \left(\prod_{j \neq i} \Phi\left((x - u(A_j))\sqrt{nq(A_j)}\right)\right) \sqrt{nq(A_i)}\phi\left((x - u(A_i))\sqrt{nq(A_i)}\right) dx$$

Here we are integrating the probability that the utility estimates for all alternatives $A_j \neq A_i$ are below the estimate for A_i, and integration is with respect to the distribution of $\hat{u}(A_i)$.

In the binary-choice case, we relied on properties of Gaussian variables to obtain a simple fixed-point equation for the RSE. The fixed-point equations for the non-binary case are considerably more involved. Whether the $n > 2$ version of the model leads to a unique RSE and whether it satisfies regularity (i.e., adding an alternative to the choice set weakly lowers the choice probabilities of all other alternatives) remain open questions.

7 RSE in Games

Once we understand that representative-sampling-based choice requires an equilibrium modeling approach even in single-agent decision problems, extending this idea to interactive decision-making is rather straightforward. Indeed, as explained in Section

2, this paper is related to the game-theoretic literature that introduced sampling-based equilibrium concepts (notably Osborne and Rubinstein (1998,2003) and Salant and Cherry (2020)). In this section we substantiate this link and describe how to extend RSE to strategic-form games. We use the Prisoner's Dilemma to illustrate how the basic force captured by RSE can have significant implications for games, especially in relation to S(K) equilibrium.

For expositional simplicity, restrict attention to symmetric, finite two-person games (but allow action sets of arbitrary size). Players' action set is A and their vNM utility function is $u : A \times A \to \mathbb{R}$. Suppose player j follows a mixed strategy $q \in \Delta(A)$, where $q(a)$ denotes the probability of playing a. Then, each action $a_i \in A$ for player i induces a lottery over her payoff v, where

$$\Pr(v \mid a_i) = \sum_{a_j \in A | u(a_i, a_j) = v} q(a_j)$$

This lottery has well-defined mean and variance, denoted $m_q(a_i)$ and $\sigma_q^2(a_i)$. Based on this pair, we can construct a well-defined normal variable $N(m_q(a_i), \sigma_q^2(a_i))$. As before, let n denote a player's total sample size.

Players use naive sampling to evaluate actions, just as they did for consumption alternatives in the basic model. Specifically, a player's estimated payoff from playing a when her opponent is playing q is defined to be

$$\hat{u}_q(a) \sim N\left(m_q(a), \frac{\sigma_q^2(a)}{nq(a)}\right) \tag{9}$$

Definition 3. *A mixed strategy $q \in \Delta(A)$ is an RSE if for every $a \in A$,*

$$q(a) = \Pr\left[\hat{u}_q(a) > \hat{u}_q(a') \text{ for every other } a' \in A\right]$$

The equilibrium definition itself is a straightforward extension of RSE to symmetric two-player games. In the spirit of the sampling-based game-theoretic solution concepts described above, it captures the idea that players base their choices on samples from the equilibrium distribution. RSE introduces an additional layer of endogeneity, in that a player's sample composition also depends on her own strategy.

The extension of RSE to games involves another modeling innovation, which is the Gaussian approximation of the payoff distribution induced by each action given the opponent's mixed strategy, as given by (9). In the basic model of Section 3, we took the Gaussian noise as given without probing its origin, as is common in the stochastic choice literature. However, when we turn to games, the payoff distribution is determined by the structure of the game, and therefore the Gaussian approximation needs to be constructed more explicitly. The Central Limit Theorem means that the approximation is good when $nq(a)$ is moderately large.

7.1 An Example: The Prisoner's Dilemma

To illustrate the extended definition of RSE, consider the following symmetric 2×2 game. The action set for each player is $\{0, 1\}$. Player i's payoff is $u_i(a_i, a_j) = a_j - ca_i$, where $c \in (0, 1)$. This is a standard specification of the Prisoner's Dilemma, where the strictly dominant action $a = 0$ corresponds to defection.

As in Section 4, our main interest here is in the contrast between the predictions of RSE and the uniform-sample case.

Proposition 7. *The Prisoner's Dilemma has a unique symmetric RSE, where the probability of playing $a = 0$ is $\Phi(c\sqrt{n})$.*

Thus, RSE uniquely predicts a positive probability of cooperation, which is below $\frac{1}{2}$ and decreases with c and n. One might think that playing a strictly dominated action with positive probability is merely a consequence of sampling error. However, we now demonstrate the crucial role that representative sampling plays in this result. Specifically, compare our analysis with the uniform-sample case: a player's estimated gain from playing $a = 0$ is

$$\hat{u}(0) - \hat{u}(1) \sim N\left(c, \frac{2r(1-r)}{n} + \frac{2r(1-r)}{n}\right) = N\left(c, \frac{4r(1-r)}{n}\right)$$

where r is the probability that the player's opponent plays $a = 0$. The equilibrium condition for this uniform-sample variant is

$$r = \Pr\left\{N\left(0, \frac{4r(1-r)}{n}\right) > -c\right\} \tag{10}$$

Remark 2. *When $nc^2 > 8$, the unique solution of (10) is $r = 1$.*

This example demonstrates once again the key role of representative sampling in two-action decision problems — specifically, its enhancement of the perceived value of inferior actions. In the Prisoner's Dilemma (as in any simultaneous-move game), the distribution of a single sample point for a player's action is given by the opponent's mixed strategy. As this strategy becomes more skewed in favor of the superior action (defection), its variance vanishes and makes the player's assessment of the two actions more accurate. Under a uniform sample, this force eliminates the possibility of cooperative play when n is not too small. The representative-sample assumption introduces a counter-force that favors the inferior action (cooperation) and therefore manages to sustain it with positive equilibrium probability for *any* value of n.

Comment. Arigapudi et al. (2021) study $S(K)$ equilibria in the Prisoner's Dilemma and their dynamic convergence properties. They show that for some range of values of K and the payoff parameters, cooperation can be part of a stable $S(K)$ equilibrium. However, if K is not small enough relative to the parameters that correspond to c in the present example, cooperation cannot be sustained in equilibrium. The uniform-sample version of the present model serves as a normal approximation of the analysis in Arigapudi et al. (2021), where $K = n/2$.

7.2 Dynamic Games: An Infinite-Horizon Trust Game

Consider the following overlapping-generations version of the Prisoner's Dilemma. Imagine time as stretching to infinity in both directions, i.e., $t = \ldots - 2, 1, 0, 1, 2, \ldots$ At every period t, a *distinct* agent, referred to as player t, chooses an action $a_t \in \{0, 1\}$. Player t's payoff is purely a function of a_t and a_{t+1}, given by $u(a_t, a_{t+1}) = a_{t+1} - ca_t$, where $c \in (0, 1)$ is a constant. This is the same payoff function as in Section 7.1, where $a_t = 1$ means that player t decides to "put her trust" in player $t + 1$.

Players have *limited recall*: they can only condition their action on the *two* most recent actions. The set of relevant truncated histories is $H = \{0, 1\}^2$. For every truncated history $h = (a_{t-2}, a_{t-1})$, (h, a_t) is a shorthand notation for the concatenated truncated history (a_{t-1}, a_t). A behavioral strategy for any player t in this game is a function $f : H \to [0, 1]$, where $f(h)$ is the probability that $a_t = 1$ given the truncated history h.

Benchmark: Nash equilibrium

As usual, this game has a Nash equilibrium in which every player chooses $a = 0$

after every history. This is the *unique* symmetric Nash equilibrium if we impose the following refinement: player t's equilibrium strategy conditions on an action in her truncated history only when she believes that this action affects the behavior of player $t + 1$.[2] The reason is as follows. Fix a candidate Nash equilibrium. Define $m^* \leq 2$ as the effective recall associated with this equilibrium. For example, $m^* = 1$ means that players condition their behavior on the most recent action, but they do not condition on earlier actions. Suppose $m^* > 0$, and consider player t's reasoning. By the definition of m^*, this player knows that player $t + 1$ will not condition her behavior on a_{t-m^*}. By the refinement, player t herself will not condition her action on a_{t-m^*}, contradicting the definition of m^*. It follows that $m^* = 0$, which means that players never condition their behavior on the history. This makes $a = 0$ the unique best-reply for each player.

The game also has symmetric Nash equilibria in which players cooperate. For instance, every f that satisfies $f(h, 1) - f(h, 0) = c$ is a symmetric Nash equilibrium, because players are always indifferent between the two actions. This equilibrium violates the criterion that players condition on a past action only when they believe it is relevant for predicting future behavior.

Let us extend the definition of RSE to this dynamic setting, where players evaluate actions at specific truncated histories. It makes sense to assume that frequently visited truncated histories will generate more observations than rarely visited ones. Therefore, a proper extension of RSE should take into account not only action frequencies, but also the (endogenous) frequencies of the truncated histories at which the actions are evaluated.

Formally, a behavioral strategy f induces a discrete-time Markov process, in which the set of states is the set of truncated histories H. The probabilities of transition from $h \in H$ into the concatenated truncated histories $(h, 1)$ and $(h, 0)$ are $f(h)$ and $1 - f(h)$, respectively. If $f(h) \in (0, 1)$ for every h — i.e., f has full support — then the Markov process is irreducible and therefore has a unique invariant distribution over H, denoted α_f. Moreover, this distribution has full support.

For every $h \in H$ and $a \in \{0, 1\}$, define the following independently distributed,

[2]This refinement is consistent with the idea that players prefer not to use complex strategies unless they have a strict benefit from doing so, as in Rubinstein (1986).

normal random variable:

$$\hat{f}(h,a) \sim N\left(f(h,a), \frac{f(h,a)(1-f(h,a))}{n\alpha_f(h,a)}\right) \tag{11}$$

This variable represents a player's estimate of the probability that the subsequent player will choose $a = 1$ following the truncated history (h, a). The origin of this expression is as follows. When a player acts at the truncated history h, she obtains a total of $n\alpha_f(h,a)$ observations about the consequences of playing a at h. Each observation is a Bernoulli distribution with success rate $f(h,a)$. Formula (11) is a normal approximation of the distribution of the sample average induced by this Bernoulli distribution.

Definition 4 (RSE in the trust game)**.** *A full-support strategy f is an RSE if, for every $h \in H$,*

$$f(h) = \Pr(\hat{f}(h,1) - \hat{f}(h,0) \geq c) \tag{12}$$

where $\hat{f}$ is defined by (11).

The following result shows that RSE involves non-stationary strategies. In particular, it implies positive reciprocity.

Proposition 8. *In any RSE, $f(a_{t-2}, a_{t-1})$ is strictly increasing in a_{t-1}.*

The message of this result is that reciprocity emerges naturally under RSE, when the representative-sample principle extends to the truncated histories at which they evaluate actions. Note also that RSE satisfies the criterion that players condition on a past action only when they believe it is relevant for predicting their opponent's behavior.

To see the logic behind the result, note that in equilibrium, one of the two actions is objectively better for all players, regardless of the history. Indeed, if $f_1 - f_0 < c$ ($> c$), defection (cooperation) is strictly better. Furthermore, the inferior action will be played with frequency below 50% after any truncated history — just as alternative A was chosen with probability below $\frac{1}{2}$ in the binary choice model. To fix ideas, assume cooperation ($a = 1$) is the inferior action. Then, since cooperation is less frequent

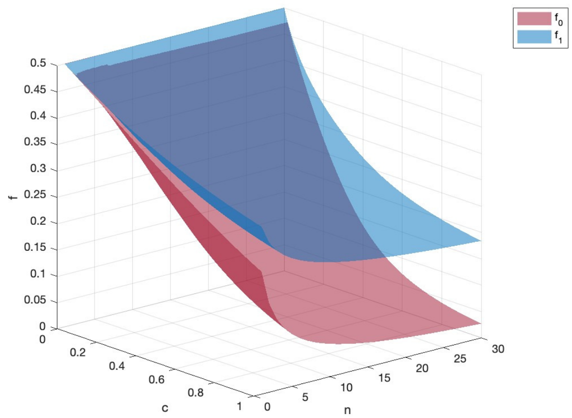

Figure 2

than defection, agents will have fewer observations about what happens after the truncated history $(h, 1)$ compared to the history $(h, 0)$. This means that their payoff estimates following $(h, 1)$ will be noisier, leading them to choose the inferior action $a = 1$ with higher probability after $(h, 1)$ than after $(h, 0)$. The same logic holds if defection is the inferior action.

While it seems plausible that $a = 1$ should be the inferior action in every equilibrium, we have been unable to prove this so far. The same holds for the question of uniqueness of RSE. Both conjectures are supported by numerical computations of RSE, the results of which are presented in Figure 2. This figure presents the values of f_0 and f_1 as a function of the parameters c and n, in the unique RSE that we could solve for numerically. Finally, whether the reciprocity patterns hold when both players have longer recall is another open question.

By incorporating the (endogenous) frequencies of histories, RSE takes us further away from the "active experimentation" image behind $S(K)$ equilibrium and brings us closer to a sampling-based equilibrium concept in which sample data is *observational* in nature.

8 Conclusion

This paper introduced a modeling innovation to the literature on stochastic choice and explored its implications in binary choice environments. We adopted a naive-sampling interpretation of the ARU model with Gaussian noise, and introduced representative sampling to capture the idea that the precision of a DM's signal about an alternative increases with its choice frequency. This, in turn, required an *equilibrium* approach to modeling single-agent stochastic choice. To facilitate an extension of these ideas to games, we used Gaussian signals as a modeling approximation, such that the mean and variance of the Gaussian signal are given by the payoff distribution induced by players' equilibrium mixed strategies.

The main economic insight that emerged from our exercise was the equilibrium force that favors inferior alternatives. This force implies very slow convergence to rational choice as sample size increases, and positive cooperation rates in the Prisoner's Dilemma for any sample size. For the same reason, convergence to rational choice is very slow as the underlying utility difference between the two alternatives rises. This means that RSE exhibits fat-tail patterns in binary stochastic choice, even though the underlying noise structure is Gaussian and therefore thin-tailed. The same equilibrium force has nuanced implications in environments with more complex information structures, as demonstrated by the comparative statics with respect to the coarseness of the DM's data, or by the reciprocity patterns in the dynamic trust game.

Finally, RSE can serve as a modeling gadget in behavioral IO or political-economy applications, where firms (political parties) compete for consumers (voters) who evaluate alternatives via naive extrapolation from finite samples.

Appendix: Proofs

As mentioned in the text, Proposition 1 is a special case of Proposition 1^*. For Remark 1, note that monotonicity in t follows as a special case of Proposition 4, and monotnicity in n can be proved with exactly the same argument since t and $\sqrt{n}$ have identical roles in the equilibrium condition in (3). Note that Propositions 2 and 3 rely on the uniqueness of RSE, i.e., on Proposition 1^*.

Proof of Proposition 2

Assume the contrary — i.e., there exists $q^* < 1$ such that for every $n > 0$, there exists $n' > n$ such that $q(n') < q^*$. Recall that $q(n') > \frac{1}{2}$. Therefore, for all such n',

$$q(n')(1 - q(n')) > q^*(1 - q^*)$$

Consequently, $\sqrt{n'q(n')(1-q(n'))}$ diverges with n', which implies that, from some point onward,

$$\Phi\left(t\sqrt{n'q(n')(1-q(n'))}\right) > q^*$$

a contradiction. ■

Proof of Proposition 3

We will prove that for all $k > 0$,

$$q(n) \leq \Phi(n^k)$$

from some $n(k)$ onward. The general claim follows immediately with a suitable change of $n(k)$. Let $n, k > 0$ and denote $x = q(n)$. That is, x is the unique solution to

$$x = \Phi\left(\sqrt{nx(1-x)}\right)$$

Assume that $x > \Phi(n^k)$. Since Φ is monotonically increasing, $\sqrt{nx(1-x)} > n^k$ or, equivalently,

$$x(1-x) > n^{2k-1} \tag{13}$$

The contradiction is immediate for $k \geq \frac{1}{2}$. Henceforth, we assume $k < \frac{1}{2}$.

Let $f(x) = x(1-x)$. The function f is invertible for $x \in [\frac{1}{2}, 1]$ with $f^{-1} : [0, \frac{1}{4}] \to [\frac{1}{2}, 1]$ given by $f^{-1}(x) = \frac{1+\sqrt{1-4x}}{2}$. The inequality (13) implies $0 < n^{2k-1} < \frac{1}{4}$ and, since f is strictly decreasing for $x \in [\frac{1}{2}$,1], also implies,

$$x < f^{-1}(n^{2k-1}) = \frac{1+\sqrt{1-4n^{2k-1}}}{2}$$

Thus,

$$\Phi(n^k) < x < \frac{1+\sqrt{1-4n^{2k-1}}}{2}$$

Hence, it suffices to show that from some n onward,

$$\Phi(n^k) \geq \frac{1+\sqrt{1-4n^{2k-1}}}{2}$$

By the Chernoff bound for the normal distribution (e.g., see Boucheron et al. (2013)),

$$1 - \Phi(x) \leq e^{-\frac{x^2}{2}} \tag{14}$$

for all $x > 0$. Thus, $\Phi(n^k) \geq 1 - e^{-\frac{n^{2k}}{2}}$ and it suffices to prove

$$e^{-\frac{n^{2k}}{2}} \leq \frac{1 - \sqrt{1 - 4n^{2k-1}}}{2} \tag{15}$$

for sufficiently large n. To see this, define

$$h(n) = \frac{1 - \sqrt{1 - 4n^{2k-1}}}{2} - e^{-\frac{n^{2k}}{2}}$$

Note that (since $k < \frac{1}{2}$) $lim_{n \to \infty} h(n) = 0$. We now show that there exists $n(k)$ such that for all $n \geq n(k)$, $h'(n) < 0$. This will imply $h(n) \geq 0$ for all $n \geq n(k)$ and thus that (15) holds for all such n. We have

$$h'(n) = \frac{(2k-1)n^{2k-2}}{\sqrt{1 - 4n^{2k-1}}} + kn^{2k-1} e^{-\frac{n^{2k}}{2}}$$

Therefore, $h'(n) < 0$ if and only if

$$\frac{e^{\frac{n^{2k}}{2}}}{n\sqrt{1 - 4n^{2k-1}}} > \frac{k}{1 - 2k}$$

Successive applications of L'Hôpital's rule imply

$$lim_{n \to \infty} \frac{e^{\frac{n^{2k}}{2}}}{n\sqrt{1 - 4n^{2k-1}}} = \infty$$

which completes the proof. ■

Proof of Proposition 1*

Equation (7) defines a fixed point of a continuous mapping from $[0, 1]^{|T|}$ to itself. Such a fixed point exists, by Brouwer's fixed-point theorem. Therefore, an RSE exists.

We prove uniqueness by contradiction. Without loss of generality, consider a trivial partition in which the only cell is T. Assume that $q = (q_t)_{t \in T}$ and $q' = (q'_t)_{t \in T}$ are both RSE solutions and $q \neq q'$. Then, there exists $t \in T$ for which $q_t \neq q'_t$. So, by (7), $\bar{q}' \neq \bar{q}$. Assume without loss of generality that $\bar{q} > \bar{q}'$. Since $t > 0$ for every $t \in T$, we have $q_t, q'_t > \frac{1}{2}$ for every $t \in T$ and hence $\bar{q} > \bar{q}' > \frac{1}{2}$. This implies

$\bar{q}(1-\bar{q}) < \bar{q}'(1-\bar{q}')$. Thus, for all $t \in T$,

$$q_t = \Phi\left(t\sqrt{n\bar{q}(1-\bar{q})}\right) < \Phi\left(t\sqrt{n\bar{q}'(1-\bar{q}')}\right) = q'_t$$

Hence,

$$\bar{q} = \sum_{t \in T} \mu_t q_t(z) < \sum_{t \in T} \mu_t q'_t(z) = \bar{q}'$$

a contradiction. ■

Proof of Proposition 4

Suppose that $\pi \succ \pi'$, and assume that $\bar{q}_{\pi'} \geq \bar{q}_{\pi}$. As we already saw, since $t > 0$ for every $t \in T$, $q_t > \frac{1}{2}$ for every t in RSE, and therefore $\bar{q}_{\pi} > \frac{1}{2}$. It follows that $\bar{q}_{\pi}(1-\bar{q}_{\pi}) \geq \bar{q}_{\pi'}(1-\bar{q}_{\pi'})$. Since $t > t'$ for every $t \in \pi$, $t' \in \pi'$, it follows from (7) that $q_t > q_{t'}$ in RSE for every $t \in \pi$, $t' \in \pi'$, hence $\bar{q}_{\pi} > \bar{q}_{\pi'}$, a contradiction. ■

Proof of Proposition 5

We prove part (i); the proof of part (ii) follows the same logic. Suppose $\bar{q}_{T^k} > \bar{q}_{T^*}$ for some $k \in \{1, ..., m\}$. Then, since both quantities are above $\frac{1}{2}$,

$$\bar{q}_{T^k}(1-\bar{q}_{T^k}) < \bar{q}_{T^*}(1-\bar{q}_{T^*})$$

By (3),

$$q_t = \Phi\left(t\sqrt{n\bar{q}_{T^*}(1-\bar{q}_{T^*})}\right)$$

for every $t \in T^*$. Therefore, since Φ is an increasing function,

$$q_t > \Phi\left(t\sqrt{n\bar{q}_{T^k}(1-\bar{q}_{T^k})}\right)$$

for every $t \in T^*$. Taking an average over $t \in T^k$ with respect to the conditional type distribution given T^k, we obtain

$$\bar{q}_{T^k} - \sum_{t \in T^k} \frac{\mu_t}{\sum_{t \in T^k} \mu_t} \Phi\left(t\sqrt{n\bar{q}_{T^k}(1-\bar{q}_{T^k})}\right) > 0 \tag{16}$$

By comparison, the definition of q' requires

$$\bar{q}'_{T^k} - \sum_{t \in T^k} \frac{\mu_t}{\sum_{t \in T^k} \mu_t} \Phi\left(t\sqrt{n\bar{q}'_{T^k}(1-\bar{q}'_{T^k})}\right) = 0 \tag{17}$$

Since the L.H.S of (16)-(17) is an increasing function of a scalar variable ($\bar{q}_{T^k}$ in the inequality, $\bar{q}'_{T^k}$ in the equation), it follows that $\bar{q}'_{T^k} < \bar{q}_{T^k}$. ■

Proof of Proposition 6

For notational simplicity only, we set $n = 1$ in what follows. Take two interval partitions Π^c and Π^f, such that Π^f is a refinement of Π^c. For notational simplicity, let $q_t^f = q_t(\Pi^f)$ and $q_t^c = q_t(\Pi^c)$.

Consider some cell $T^* \in \Pi^c$. Denote

$$\alpha_t = \frac{\mu_t}{\sum_{s \in T^*} \mu_s}$$

Define

$$Q^c = \sum_{t \in T^*} \alpha_t q_t^c = \sum_{t \in T^*} \alpha_t \Phi\left(t\sqrt{Q^c\left(1 - Q^c\right)}\right)$$

This is the average equilibrium probability of choosing B among types in T^* under the partition Π^c.

Obviously, if T^* is also a cell in Π^f, then $q_t^c = q_t^f$ for every $t \in T^*$, hence $Q^C = Q^f$. We now turn to the non-degenerate case, in which Π^f strictly refines the cell T^*. Let β_π be the probability of $\pi \in \Pi^f$ conditional on $\pi \subset T^*$. Denote

$$\bar{q}_\pi = \sum_{s \in \pi} \frac{\alpha_s}{\beta_\pi} q_s^f$$

Define

$$Q^f = \sum_{t \in T^*} \alpha_t q_t^f = \sum_{t \in T^*} \alpha_t \Phi\left(t\sqrt{\bar{q}_{\Pi^f(t)}(1 - \bar{q}_{\Pi^f(t)})}\right)$$

This is the equilibrium probability of choosing B conditional on $t \in T^*$ under Π^f. Suppose that $Q^c \leq Q^f$. Then, since $\sqrt{q(1-q)}$ is strictly decreasing in $q > \frac{1}{2}$,

$$\sqrt{Q^c(1 - Q^c)} \geq \sqrt{Q^f(1 - Q^f)}$$

Since Φ is strictly increasing,

$$Q^c = \sum_{t \in T^*} \alpha_t \Phi\left(t\sqrt{Q^c\left(1 - Q^c\right)}\right) \geq \sum_{t \in T^*} \alpha_t \Phi\left(t\sqrt{Q^f\left(1 - Q^f\right)}\right)$$

Denote

$$x_\pi = \sqrt{\bar{q}_\pi(1-\bar{q}_\pi)}$$

The expression $\sqrt{q(1-q)}$ is strictly concave in q. Therefore,

$$\begin{aligned}\sqrt{Q^f(1-Q^f)} &= \sqrt{\left(\sum_{\pi\subset T^*}\beta_\pi\bar{q}_\pi\right)\left(1-\sum_{\pi\subset T^*}\beta_\pi\bar{q}_\pi\right)} \\ &> \sum_{\pi\subset T^*}\beta_\pi\sqrt{\bar{q}_\pi(1-\bar{q}_\pi)} = \sum_{\pi\subset T^*}\beta_\pi x_\pi\end{aligned}$$

Define the function $H(s,x) = \Phi(sx)$ where $s, x > 0$. Since Φ is strictly increasing,

$$\sum_{t\in T^*}\alpha_t\Phi(t\sqrt{Q^f(1-Q^f)}) > \sum_{t\in T^*}\alpha_t\Phi\left(t\sum_{\pi\subset T^*}\beta_\pi x_\pi\right) = \sum_{t\in T^*}\alpha_t H\left(t, \sum_{\pi\subset T^*}\beta_\pi x_\pi\right)$$

By concavity of H with respect to its second argument,

$$H\left(t, \sum_{\pi\subset T^*}\beta_\pi x_\pi\right) > \sum_{\pi\subset T^*}\beta_\pi H(t, x_\pi)$$

for every t. Therefore,

$$\sum_{t\in T^*}\alpha_t H\left(t, \sum_{\pi\subset T^*}\beta_\pi x_\pi\right) > \sum_{t\in T^*}\sum_{\pi\subset T^*}\alpha_t\beta_\pi H(t, x_\pi)$$

Note that $x_\pi \in (0, \frac{1}{2})$ for every π, by the definition of x_π. Furthermore, by Proposition 4, the cells in Π^f are ordered such that $\bar{q}_{\Pi^f(t)}$ is increasing in t, and hence $x_{\Pi^f(t)}$ is decreasing in t. By Lemma 1 below, H is supermodular when $t < 2$. Therefore,

$$\begin{aligned}\sum_{t\in T^*}\sum_{\pi\subset T^*}\alpha_t\beta_\pi H(t, x_\pi) &> \sum_{t\in T^*}\alpha_t H(t, x_{\Pi^f(t)}) \\ &= \sum_{t\in T^*}\alpha_t\Phi\left(t\sqrt{\bar{q}_{\Pi^f(t)}(1-\bar{q}_{\Pi^f(t)})}\right) = Q^f\end{aligned}$$

This inequality is a special case of a standard inequality from the literature on stochastic orderings — e.g., see Tchen (1980).[3] We have thus obtained $Q^c > Q^f$, a contradiction. It follows that for every cell $T^* \in \Pi^c$, $Q^c \geq Q^f$, with a strict inequality for at least one cell. Therefore, $\bar{q}(\Pi^c) > \bar{q}(\Pi^f)$. ■

Lemma 1. *Let $H(s,x) = \Phi(sx)$ where $s, x > 0$. If $s < 2$ and $x \in (0, \frac{1}{2})$, then H is supermodular.*

Proof of Lemma 1

Recall that

$$H(s,x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{sx} e^{-\frac{a^2}{2}} da$$

The cross derivative of H is

$$\frac{\partial H(s,x)}{\partial x \partial s} = \frac{e^{-\frac{(xs)^2}{2}}}{\sqrt{2\pi}} \left[1 - (xs)^2\right]$$

When $x < \frac{1}{2}$, this expression is strictly positive whenever $s < 2$. ■

Proof of Proposition 7

Let q denote the RSE probability of $a = 0$. When a player draws a single sample point from an action a, she obtains the payoff $1 - ca$ with probability $1 - q$ and the payoff $-ca$ with probability q. The normal distribution that shares the mean and variance with this random variable is

$$N(1 - q - ca, q(1-q))$$

In RSE, the player samples $a = 0$ nq times and $a = 1$ $n(1-q)$ times. Therefore, the player's estimated gain from playing $a = 0$ is

$$\hat{u}(0) - \hat{u}(1) \sim N\left(c, \frac{q(1-q)}{nq} + \frac{q(1-q)}{n(1-q)}\right) = N\left(c, \frac{1}{n}\right)$$

In RSE,

$$q = \Pr\left\{N\left(0, \frac{1}{n}\right) > -c\right\} = \Phi(c\sqrt{n})$$

This completes the proof. ■

[3]We thank Meg Meyer for the reference.

Proof of Remark 2
The condition (10) can be rewritten as

$$r = \Phi\left(c\sqrt{\frac{n}{4r(1-r)}}\right)$$

Applying the Chernoff bound (14), we obtain

$$r = \Phi\left(c\sqrt{\frac{n}{4r(1-r)}}\right) \geq 1 - e^{-\frac{c^2 n}{8r(1-r)}}$$

This inequality is equivalent to

$$x \leq e^{-\frac{c^2 n}{8x(1-x)}}$$

where $x = 1 - r$. We now show that when $nc^2 > 8$, this inequality fails for all $x \in (0, 1]$. To see this, denote $t = c^2 n$ and define

$$f(x,t) = x - e^{-\frac{t}{8x(1-x)}}$$

Note that for all $x > 0$, $f(x,t)$ is increasing in t for $t > 0$. Thus, it suffices to prove that $f(x, 8) > 0$ for all $x \in (0, 1]$. For all such x we have $x > x(1-x) > 0$ and hence,

$$f(x,8) = x - e^{-\frac{1}{x(1-x)}} > x - e^{-\frac{1}{x}}$$

The R.H.S can easily be shown to be strictly positive for all $x > 0$. ■

Proof of Proposition 8

First, observe that $f(h)$ only depends on the most recent action, i.e., $f(h) = f(a_{t-2}, a_{t-1})$ is constant in a_{t-2}. Indeed, by equations (11) and (12), $f(h)$ is pinned down by $f(h,1), f(h,0), \alpha_f(h,0)$ and $\alpha_f(h,1)$, which by definition do not depend on the earliest action in the truncated history h (e.g., if $h = (a_{t-2}, a_{t-1})$ then $(h,a) = (a_{t-1}, a)$).

Accordingly, we denote by f_a the probability that $a_{t+1} = 1$ conditional on $a_t = a$. In a similar vein, we use the notation α_h for $\alpha_f(h)$. Then, condition (12) can be written as

$$f_1 = \Pr\left(\hat{f}(1,1) - \hat{f}(1,0) \geq c\right)$$
$$f_0 = \Pr\left(\hat{f}(0,1) - \hat{f}(0,0) \geq c\right)$$

where

$$\hat{f}(1,1) - \hat{f}(1,0) \sim N\left(f_1 - f_0, \frac{f_1(1-f_1)}{n\alpha_{11}} + \frac{f_0(1-f_0)}{n\alpha_{10}}\right)$$
$$\hat{f}(0,1) - \hat{f}(0,0) \sim N\left(f_1 - f_0, \frac{f_1(1-f_1)}{n\alpha_{01}} + \frac{f_0(1-f_0)}{n\alpha_{00}}\right)$$

By the definition of α_f,

$$\begin{aligned}
\alpha_{11} &= f_1 \cdot (\alpha_{11} + \alpha_{01}) \\
\alpha_{10} &= (1 - f_1) \cdot (\alpha_{11} + \alpha_{01}) \\
\alpha_{01} &= f_0 \cdot (\alpha_{10} + \alpha_{00}) \\
\alpha_{00} &= (1 - f_0) \cdot (\alpha_{10} + \alpha_{00}) \\
1 &= \alpha_{00} + \alpha_{01} + \alpha_{10} + \alpha_{11}
\end{aligned}$$

The solution for α_f is

$$\begin{aligned}
\alpha_{11} &= \tfrac{f_1 f_0}{1+f_0-f_1} \\
\alpha_{10} &= \tfrac{f_0(1-f_1)}{1+f_0-f_1} \\
\alpha_{01} &= \tfrac{f_0(1-f_1)}{1+f_0-f_1} \\
\alpha_{00} &= \tfrac{(1-f_0)(1-f_1)}{1+f_0-f_1}
\end{aligned}$$

If $f_1 - f_0 \geq c$, then $f_1 > f_0$ and we are done. Now suppose $f_1 - f_0 < c$. Then, $f_1, f_0 < \frac{1}{2}$. Therefore, $\alpha_{11} < \alpha_{01}$ and $\alpha_{10} < \alpha_{00}$. It follows that $\hat{f}(1,1) - \hat{f}(1,0)$ and $\hat{f}(0,1) - \hat{f}(0,0)$ have the same mean, and

$$Var(\hat{f}(0,1) - \hat{f}(0,0)) < Var(\hat{f}(1,1) - \hat{f}(1,0))$$

Since the mean lies below c,

$$f_1 = \Pr(\hat{f}(1,1) - \hat{f}(1,0) \geq c) > \Pr(\hat{f}(0,1) - \hat{f}(0,0) \geq c) = f_0 \quad \blacksquare$$

References

Arigapudi, S., Y. Heller and I. Milchtaich (2021), Instability of defection in the prisoner's dilemma under best experienced payoff dynamic, Journal of Economic Theory 197, 105174.

Banerjee, A. (1992), A simple model of herd behavior, Quarterly Journal of Economics, 107(3), 797-817.

Banerjee, A. and D. Fudenberg (2004), Word-of-mouth learning, Games and Economic Behavior 46, 1-22.

Bikhchandani, S., D. Hirshleifer and I. Welch (1992), A theory of fads, fashion, custom, and cultural change as informational cascades, Journal of Political Economy 100, 992-1026.

Boucheron, S., G. Lugosi and P. Massart (2013), Concentration inequalities: A nonasymptotic theory of independence, Oxford university press.

De Langhe, B., P. Fernbach and D. Lichtenstein (2016), Navigating by the stars: Investigating the actual and perceived validity of online user ratings, Journal of Consumer Research 42, 817-833.

Ellison, G. and D. Fudenberg (1995), Word-of-mouth communication and social learning, Quarterly Journal of Economics 110, 93-125.

Fudenberg, D., G. Lanzani and P. Strack (2024), Limited memory, learning, and stochastic choice, mimeo.

Golub, B. and M. Jackson (2012), How homophily affects the speed of learning and best-response dynamics, Quarterly Journal of Economics 127, 1287-1338.

Goncalves, D. (2020), Sequential sampling and equilibrium, mimeo.

Jehiel, P. (2005), Analogy-based expectation equilibrium, Journal of Economic Theory 123, 81-104.

Lu, J. (2016). Random choice and private information, Econometrica 84(6), 1983-2027.

Obrecht, N., G. Chapman and R. Gelman (2007), Intuitive t tests: Lay use of statistical information, Psychonomic Bulletin & Review 14, 1147-1152.

Osborne, M. and A. Rubinstein (1998), Games with procedurally rational players, American Economic Review, 834-847.

Osborne, M. and A. Rubinstein (2003), Sampling equilibrium with an application to strategic voting, Games and Economic Behavior 45, 434-441.

Natenzon, P. (2019), Random choice and learning, Journal of Political Economy 127, 419-457.

Salant, Y. and J. Cherry (2020), Statistical inference in games, Econometrica 88, 1725-1752.

Sethi, R. (2000), Stability of equilibria in games with procedurally rational players, Games and Economic Behavior, 32, 85-104.

Shapiro, C. (1983), Optimal pricing of experience goods, The Bell Journal of Economics, 497-507.

Spiegler, R. (2006a), The market for quacks, Review of Economic Studies 73, 1113-1131.

Spiegler, R. (2006b), Competition over agents with boundedly rational expectations, Theoretical Economics, 1, 207-231.

Spiegler, R. (2011), Bounded Rationality and Industrial Organization, Oxford University Press.

Strzalecki, T. (2023), Stochastic Choice Theory, Preprint.

Szech, N. (2011), Becoming a bad doctor, Journal of Economic Behavior and Organization 80, 244-257.

Tchen, A. (1980), Inequalities for distributions with given marginals, Annals of Probability 8, 814-827.

Thurstone, L. L. (1927), A law of comparative judgment, Psychological review, 34(4), 273.

Tversky, A. and D. Kahneman (1971), Belief in the law of small numbers, Psychological Bulletin 76, 105-110.