A Representative-Sampling Model of Stochastic Choice*

Tuval Danenberg[†] and Ran Spiegler[‡]

October 4, 2023

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 the Foerder Institute and the Sapir Center is gratefully acknowledged. We thank Jose Apesteguia, Ian Ball, Drew Fudenberg, Nathan Hancart, Paola Manzini, Meg Meyer, Yuval Salant, Tomasz Strzalecki, and seminar participants at NHH and MIT for helpful feedback.

[†]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 $u(a)$. However, 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 noisy outcome of an introspective process by which the DM tries to access the intrinsic value of a choice alternative. Then, when the DM consumes an alternative more frequently, she is likely to have an easier time retrieving memories of consumption experiences. Now 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 binary-choice 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 payoff $u(z) + \varepsilon$, where ε is an independent draw from a *normal distribution* with mean zero and variance σ^2. Thus, the DM's Gaussian signal 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 payoff 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 deciding between red or white wine for dinner. Suppose the agent would derive greater expected pleasure from drinking white wine. However, she does not know her taste for wine well enough to recognize this. Instead, she relies on her personal memory of previous wine experiences. These experiences are noisy due to variations in grape type, vintage, or dish pairings. Importantly, the composition of the sample will reflect the agent's previous choices: if she tended to drink white wine in the past, she will have a more precise understanding of her pleasure from this type of wine. 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 white wine 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 difficulties 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 wine example, the assumption that red wine is intrinsically inferior (according to the DM's true underlying taste) leads the DM to have fewer sample points about red wine, which makes her assessment of red wine noisier. Since a noisy assessment favors an intrinsically inferior alternative, we have an equilibrium effect that magnifies the choice frequencies of inferior alternatives. After establishing existence, uniqueness, and monotonicity results for RSE, our main result addresses the implications of this basic insight for how choice frequencies depend on the sample size n. The equilibrium force described above implies that not only does representative sampling increase the equilibrium frequency of the inferior alternative relative to the rational or uniform-sample benchmarks, but the *rate* with which this frequency vanishes with n is extremely slow.

We also consider an extension of this binary-choice model, in which the DM has multiple types, defined by their intrinsic utility difference between the two alternatives (all types agree on the sign of this difference). The types are partitioned into "intervals", such that each type's sample is restricted to the interval that includes it. This extension captures cases in which an agent in a given situation only accesses data about *similar* situations. We carry out comparative statics with respect to the coarseness of the partition. In particular, we show that when the objective payoff difference between the two alternatives is not too large, a finer partition leads to a *higher* overall equilibrium probability of choosing the intrinsically inferior alternative. In terms of the wine example, if the DM shifts from a broad consideration of all wines to more specific comparisons — e.g., French reds against French whites — their likelihood of opting for the lesser-quality red wine increases, even when the total sample size for each decision is held fixed.

Finally, we suggest how RSE can be extended from decision problems to games. We illustrate this direction with the Prisoner's Dilemma, and show that unlike the rational and uniform-sample benchmarks, the favoring-inferior-alternatives effect of RSE leads to positive cooperation rates in equilibrium.

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. For example, in Strzalecki (2023), the learning interpretation of random choice focuses on dynamic sampling procedures that are more consistent with Bayesian rationality.

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) study 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) studied 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.[1]

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)).[2]

The physical-sampling interpretation of our model links it to the literatures on word-of-mouth learning (e.g., Ellison and Fudenberg (1995) or 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

[1] For a sampling-based game-theoretic solution concept that adheres to Bayesian rationality, see Goncalves (2020).

[2] The observation that frequently used products lead 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.

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.

2 Model

A DM faces a choice between two alternatives, denoted A and B. The DM's type is $t \in T$, where $T \subset \mathbb{R}$ is a finite set. Let $\mu \in \Delta(T)$ represent a distribution over types in a large population of agents facing the same choice problem. Denote the share of type t in the population by μ_t. The DM's objective expected payoff from choosing an alternative $z \in \{A, B\}$ given her type $t \in T$ is $u(z,t)$.

Let $q_t(z)$ be the probability a DM of type t chooses z. The average frequency of choosing z in the population is

$$\bar{q}(z) = \sum_{t \in T} \mu_t q_t(z) \tag{1}$$

We will often use the abbreviated notation $q_t = q_t(B)$ and $\bar{q} = \bar{q}(B)$.

In our model, q_t is a consequence of agents' attempt to learn their payoffs from samples. A 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:

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

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

Definition 1 *A profile $(q_t)_{t \in T}$ is a* ***representative-sampling equilibrium (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 (2).

The idea behind this formulation is as follows. Before choosing an action, a DM of type t samples the payoff realizations of each alternative. The alternatives' representation in her sample matches their choice frequencies among the types in the population. The DM is a "naive frequentist", taking sample outcomes at face value. That is, she regards the sample average $\hat{u}(z,t)$ as an accurate representation of her underlying average payoff from choosing z, ignoring sampling error.

By the assumption that $\hat{u}(A,t)$ and $\hat{u}(B,t)$ are independent normal variables,

$$\hat{u}(B,t) - \hat{u}(A,t) \sim N\left(u(B,t) - u(A,t), \frac{\sigma^2}{n\bar{q}(A)\bar{q}(B)}\right) \tag{3}$$

Therefore, we can identify t with the mean of this distribution — i.e.,

$$t = u(B,t) - u(A,t)$$

such that t measures the DM's underlying intrinsic preference for B over A. Furthermore, it is clear from (3) that the value of σ can be *normalized* to 1 without loss of generality (because we can rescale t). From now on, we set $\sigma = 1$. Consequently, the equilibrium condition can be rewritten as

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

for all t, or, equivalently,

$$q_t = \Phi\left(t\sqrt{n\bar{q}(1-\bar{q})}\right) \tag{4}$$

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

We will use (4) as our working definition of RSE in Section 3. This definition immediately implies that in any RSE, a DM of type t chooses her objectively superior alternative (i.e., the z with the higher $u(z,t)$) 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.

When $\bar{q}(z) = 0$, $\hat{u}(z,t)$ is ill-defined because it involves infinite variance. To handle this, we treat $N(0,\infty)$ as a well-defined distribution satisfying $\Pr(x \leq c) = \frac{1}{2}$ for every c. Consequently, the definition of RSE given by (4) is legitimate even when $\bar{q}(z) = 0$. Equilibrium choice probabilities will always be interior.

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 average choice distribution is skewed (i.e., when $\bar{q}$ is close to zero or one), the variance of $\hat{u}(B,t) - \hat{u}(A,t)$ is large, and this introduces an equilibrium counterforce toward a less skewed distribution, namely larger choice errors. Section 3 will explore the implications of this force.

Comment: What does the DM observe?
Recall that one interpretation of $\hat{u}(z,t)$ is that it represents naive-frequentist inference from a 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 payoff 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 payoff from that same experience would be.

3 Analysis

We begin our analysis with a few elementary results.

Remark 1 *An RSE exists.*

Remark 2 *Let q be an RSE. If $t' > t$, then, $q_{t'} > q_t$.*

Both results are immediate consequences of (4). This equation 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. Furthermore, fixing an equilibrium q, the R.H.S of (4) is strictly increasing in t, and therefore q_t must increase in t.

The following result establishes equilibrium uniqueness when B is the intrinsically superior alternative for all DM types.

Proposition 1 *Assume $t > 0$ for every $t \in T$. Then, there is a unique RSE.*

Finding conditions for equilibrium uniqueness when the sign of t is not constant is an open problem.

3.1 Convergence Properties

Consider the case of a single DM type — i.e., $T \equiv \{t\}$. Let $t > 0$, without loss of generality. In this sub-section, we will use $q_t(n)$ to denote the RSE for type t and sample size n, in order to highlight the role of n. It is uniquely given by

$$q_t(n) = \Phi\left(t\sqrt{nq_t(n)(1-q_t(n))}\right) \tag{5}$$

Our task is to analyze the dependence of $q_t(n)$ on n, especially in comparison with uniform sampling.

First, observe that $q_t(n)$ increases with n, by the same logic as Remark 2. The next result shows that choice errors vanish as n tends to infinity.

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

Now compare (4) with the case of a *uniform sample*, where each alternative is sampled $\frac{n}{2}$ 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

$$r_t = \Phi\left(\frac{t}{2}n^{\frac{1}{2}}\right) \tag{6}$$

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

Formula (6) has two noteworthy features. First, it lacks the equilibrium effect that arises from representative sampling. Second, since $\sqrt{q(1-q)} < \frac{1}{2}$ for any $q \in (0,1)$, r_t assigns higher probability to the favored alternative than any RSE value of q_t, for any type t.

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

Proposition 3 *For every $k > 0$, there exists $n(k)$ such that for every integer $n \geq n(k)$: $q_1(n) \leq \Phi(\frac{1}{2}n^k)$*

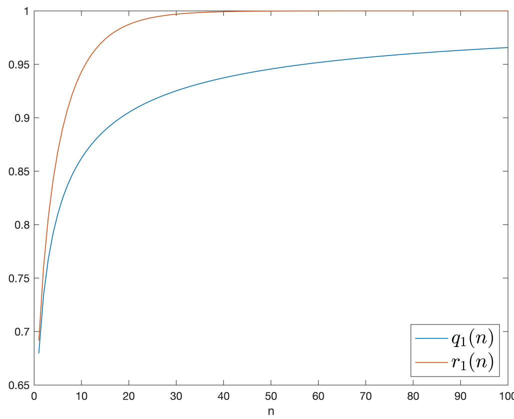

Figure 1(a)

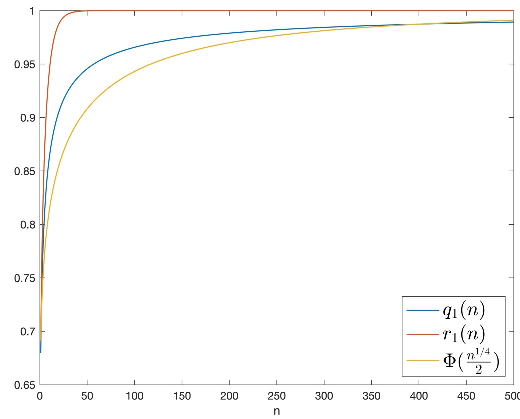

Figure 1(b)

In the uniform-sample case, $r_t(n)$ increases with n like $\Phi(\sqrt{n})$. By comparison, in the representative sample case, $q_t(n)$ 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 describes $\Phi(\frac{1}{2}n^{1/4})$). As we can see, the uniform-case specification exhibits fast convergence — e.g., $r_1(30) \approx 0.997$. In contrast, the RSE prediction is $q_1(30) \approx 0.925$. Considering that $t = 1$ represents an objective payoff difference of one standard deviation (recall that $\sigma = 1$), this is a significant choice error. Moreover, convergence is very slow such that from around $n = 400$, $q_1(n) < \Phi(\frac{1}{2}n^{1/4})$.

3.2 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 the entire population, but rather from a sub-population of "similar" agents. To capture this, we introduce a new primitive into our model, in the spirit of Jehiel's (2005) notion of analogy partitions. Let Π be a partition of T, where $\Pi(t)$ denotes the partition cell that includes t. For some of our results, we will assume that Π consists of "intervals" — i.e., if $\Pi(t) = \Pi(t')$ and $t < t'' < t'$, then $\Pi(t'') = \Pi(t)$. In this case, we refer to Π as an *interval partition*.

The average frequency of choosing z among types in $\Pi(t)$ is

$$\bar{q}_{\Pi(t)}(z) = \frac{\sum_{t \in \Pi(t)} \mu_t q_t(z)}{\sum_{t \in \Pi(t)} \mu_t} \tag{7}$$

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 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. In the introspective take, the DM accesses similar situations she experienced in the past. For these examples, a fine partition means the DM only considers situations who are very similar their own. An alternative interpretation relevant for physical sampling is that Π represents a particular word-of-mouth learning environment. The DM learns from the experiences of socially linked agents. The partition corresponds to a particular social network that consists of isolated cliques. When Π is an interval partition, a finer partition corresponds to a larger degree of homophily.

The next result establishes monotonicity of $\bar{q}_\pi$ when Π is an interval partition. Given any two cells $\pi, \pi' \in \Pi$, write $\pi' \succ \pi$ if and only if $t' > t$ for every $t \in \pi$, $t' \in \pi'$.

Proposition 4 *Suppose Π is an interval partition. Then, in equilibrium, $\pi \succ \pi'$ implies $\bar{q}_\pi > \bar{q}_{\pi'}$.*

Note that the monotonicity result applies to average choice probabilities in cells of the interval partition Π, but not necessarily to choice probabilities of individual types. 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 payoff 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. Of course, within a given cell, q_t increases with t, as in Remark 2.

We now turn to the question of how the coarseness of the partition 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. If we think of each cell in the refined partition as a "peer group", then the message of the result is that increased homophily (i.e., greater tendency to learn from similar types) brings the choice probabilities in extreme cells closer together. 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.

While Proposition 5 holds for any partitional structure, in the remainder of the subsection we restrict attention to interval partitions. Our next result will make use of the following lemma. Define the function $H(s, x) = \Phi(sx)$ where $s, x > 0$.

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

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 underlying payoff advantage of alternative B is sufficiently small, a finer partition leads to a higher probability of choice mistakes. Recall our two alternative interpretations of Π. Under the "coarse data" interpretation, the result means that finer data has an adverse effect on average choice quality. Under the "homophily" interpretation, the result means that increasing the homophily of the underlying social network that agents rely on for learning leads to poorer choice on average. 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.

4 Extension to Simultaneous-Move 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 the Introduction, 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.

For expositional simplicity, restrict attention to symmetric, finite two-person games. 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.

A player's estimated payoff from playing a when her opponent is playing q is defined to be

$$\hat{u}_q(a) \sim N\left(m_q(a), \frac{\sigma_q^2(a)}{nq(a)}\right) \tag{8}$$

Definition 2 *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 (8). In the basic model of Section 2, 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 when $nq(a)$ is moderately large, the approximation is precise.

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 $c < 1$. This is a standard specification of the Prisoner's Dilemma, where the strictly dominant action $a = 0$ corresponds to defection.

As in Section 3, 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

$$r = \Pr\left\{N\left(0, \frac{4r(1-r)}{n}\right) > -c\right\} \tag{9}$$

Remark 3 *When $nc^2 > 8$, the unique solution of (9) 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 objectively 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 objectively 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 objectively 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 $K = n/2$.

5 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 the value of 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 more complex environments such as strategic games, we introduced Gaussian signals as a modeling approximation, such that the mean and variance of the Gaussian signal are given by the objective payoff distribution induced by players' equilibrium mixed strategies.

The main economic insight that emerged was the equilibrium force that favors inferior alternatives — leading to very slow convergence to rational choice as sample size increases, and to positive cooperation rates in the Prisoner's Dilemma for any sample size. This force can have nuanced implications when coupled with varying information structures, as demonstrated by the comparative statics with respect to the coarseness of the DM's information.

References

[1] 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.

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

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

[4] 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.

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

[6] 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.

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

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

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

[10] Jehiel, P. (2005), Analogy-based expectation equilibrium, Journal of Economic theory 123, 81-104.

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

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

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

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

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

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

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

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

[19] Strzalecki, T. (2023), Stochastic Choice Theory, Preprint.

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

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

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

Appendix: Proofs

Proposition 1

Assume towards contradiction that $q = (q_t)_{t\in T}$ and $q' = (q'_t)_{t\in T}$ are both RSE solutions and $q \neq q'$. Let t satisfy $q_t \neq q'_t$ for some $t \in T$. Then, by (4), $\bar{q}' \neq \bar{q}$. 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.

Proposition 2

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

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

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

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

a contradiction.

Proposition 3

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

$$q_1(n) \leq \Phi(n^k)$$

from some $n(k)$ onward. The general claim follows immediately with a suitable change of $n(k)$. Let $n, k > 0$ and denote $x = q_1(n)$. That is, x is the unique solution to

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

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

$$x(1-x) > n^{2k-1} \tag{10}$$

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 (10) implies $0 < n^{2k-1} < \frac{1}{4}$ and, since f is strictly decreasing, 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)),

$$1 - \Phi(x) \leq e^{-\frac{x^2}{2}} \tag{11}$$

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

$$e^{-\frac{n^{2k}}{2}} \leq \frac{1-\sqrt{1-4n^{2k-1}}}{2} \tag{12}$$

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 (12) 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.

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 (4) that $q_t > q_{t'}$ in RSE for every $t \in \pi$, $t' \in \pi'$, hence $\bar{q}_{\pi} > \bar{q}_{\pi'}$, a contradiction.

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 $k = 1, ..., m$. 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 (4),

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

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

$$\bar{q}_{T^k} - \sum_{t\in T^k}\frac{\mu_t}{\sum_{t\in T^k}\mu_t}\Phi\left(t\sqrt{n\bar{q}_{T^k}(1-\bar{q}_{T^k})}\right) > 0 \tag{13}$$

By comparison, the definition of q' requires

$$\bar{q}'_{T^k} - \sum_{t\in T^k}\frac{\mu_t}{\sum_{t\in T^k}\mu_t}\Phi\left(t\sqrt{n\bar{q}'_{T^k}(1-\bar{q}'_{T^k})}\right) = 0 \tag{14}$$

Since the L.H.S of (13)-(14) 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 $\bar{q}'_{T^k} < \bar{q}_{T^k}$.

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 $s<2$.

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,

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

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 the monotonicity result, 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, 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 $T^* \in \Pi^c$, $Q^c \leq Q^f$, with a strict inequality for at least one cell. Therefore, $\bar{q}(\Pi^c) < \bar{q}(\Pi^f)$.

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

$$N\left(1 - q - ca, q(1-q)\right)$$

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.

Remark 3

The condition (9) can be rewritten as

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

Applying the Chernoff bound (11), 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

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

[3] We thank Meg Meyer for the reference.

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 $x > 0$.