Heterogeneous Agent Quantal Response Equilibrium

Jeevant Rampal Fernando Stragliotto*

April 25, 2023

Abstract

In this paper we study a setting where players of a sequential-move game may have heterogeneous skill. Skill is captured by payoff responsiveness in quantal response models. Mckelvey and Palfrey (1998) provide the Quantal Response Equilibrium for extensive-form games (AQRE) where all players are assumed to have homogeneous skill. In this paper we extend the AQRE by modeling heterogeneous skill and uncertainty and belief-updating (BU) about opponents' skills. First, we provide an equilibrium model incorporating skill-heterogeneity and uncertainty but not BU—this is called Heterogeneous AQRE (or HAQRE). Next, we incorporate naive disequilibrium belief-updating (BU) to define the HAQRE-BU. We show that these concepts exist, and in the context of finite perfect information games, they are unique, and they yield simple data applicability without fixed-point calculations. We use experimental data from a sequential-move game where players with different experience-levels interacted (Rampal (2020)) to show that modeling heterogeneity and belief updating about skills can each yield better data-fit in such settings.

Keywords: Quantal response; Sequential-move games; Heterogeneous skill; Belief updating.
JEL: D91, C72, D83

1 Introduction

The Quantal Response Equilibrium (QRE) by McKelvey and Palfrey (1995) has had wide applicability in organizing behavioral data from simultaneous-move games (see Goeree, Holt, and Palfrey (2016) for a survey). QRE models a scenario where players noisily best respond. Players have a positive chance of playing every available strategy, not just the best

*J. Rampal: Associate Professor of Economics, IIM Ahmedabad, Gujarat 380015, India. Email: jeevantr@iima.ac.in. F. Stragliotto (corresponding author): PhD student, Queen Mary University of London, Mile End Road, London E1 4NS, United Kingdom. Email: f.stragliotto@hss22.qmul.ac.uk. J. Rampal gratefully acknowledges research funding from R&P office and NSE CBS, IIM Ahmedabad. We thank Varun Bansal for his research assistance. We thank seminar audiences at BREW 2022 for their comments.

responses, and they choose each strategy with a probability proportion that is increasing in the expected payoff from that strategy, ceteris paribus. Given that QRE is an equilibrium approach, these expected payoffs account for the fact that the opponents are also playing noisy best responses. The skill of a player is captured by the player's lack-of-noise in best responding, or the player's payoff-responsiveness. The agent normal-form version of the Quantal Response Equilibrium (AQRE for short) by Mckelvey and Palfrey (1998) extended QRE to extensive-form games. The AQRE has also been applied to several sequential games (e.g., Ivanov, Levin, and Peck (2009); Tingley and Wang (2010)).

Both QRE and AQRE assume a homogeneous skill level across players. But, skills can be heterogeneous across players for a variety of reasons, for example, experience (Rampal (2020)), or cognitive differences in general (Camerer, Ho, and Chong (2004)). Rogers, Palfrey, and Camerer (2009) provide the Heterogeneous QRE (or HQRE) which models heterogeneity and uncertainty regarding skill among the players within the QRE setting of simultaneous-move games. The HQRE has provided insights regarding several scenarios like auctions (Camerer, Nunnari, and Palfrey (2016)) and investment bubbles (Moinas and Pouget (2013)). Studies analyzing capacity allocation games (Chen, Su, and Zhao (2012); Cui and Zhang (2018)) have also argued for the applicability of HQRE.

We attempt to make an analogous contribution for sequential-move games as Rogers, Palfrey, and Camerer (2009) did for simultaneous-move games.[1] In this paper, we restrict attention to finite perfect-information games and finitely repeated games. We provide a model where players of such sequential-move games can have heterogeneity in skill, and they are uncertain about each opponent's skill level. We call this the HAQRE. As we are dealing with sequential-move games, we also model belief updating (BU) about the opponent's skill within a single play of the game, a feature absent from simultaneous-move games. Finally, among these theoretical models, we study those pertaining to perfect-information games using experimental data from Rampal (2020), which tested such games among players with heterogeneous experience-levels. We find that both features, heterogeneity and belief updating yield significant improvements in data-fit relative to their respective baselines.

The importance or need for incorporating heterogeneity in the AQRE model has been directly mentioned by several studies in the literature. Rogers, Palfrey, and Camerer (2009) state: "Another important area for research is extension of these ideas (heterogeneity) to extensive form games" (page 1462). Ivanov, Levin, and Peck (2009) study an extensive-form investment game and find that their classification of subjects as following certain behavioral rules-of-thumb fits the data better than the AQRE. But, they mention that "a heterogeneous version of (A)QRE might provide a better fit in our experiment than the symmetric QRE"

[1] Rogers et al. (2009) also provide a theoretical framework to understand the links between the Level-k and quantal response models. In this paper we do not make an analogous contribution on that dimension for sequential-move games.

(footnote 37, page 1499).

Despite the need for incorporating heterogeneity outlined by the literature, in the AQRE setting "extending hierarchical or recursive approaches to extensive form games is less straightforward" (Rogers, Palfrey, and Camerer (2009), page 1462). One key challenge is that one cannot use the agent-normal-form approach as utilized to extend QRE to extensive-form games. The agent-normal-form approach is to consider each player at each move a different player, but when each player can be one of several skill-types, the agent-normal-form seems unsuitable. This is because if a player Ann can be one of two types (high or low skill), then for any skill realization, she should be the same skill-type at all her moves. By contrast, the agent-normal-form approach would consider Ann at stage-1 a different player from Ann at stage-3. Furthermore, to incorporate belief updating meaningfully, each opponent of Ann should account for the fact that the same player, Ann, is moving at all her moves.

Our modeling approach can be illustrated by how we treat perfect information games. Without belief updating, HAQRE is simple. We start with a finite perfect-information game and assume that the "true game" is a distorted version of the perfect-information game: in "truth" each player can have different levels of skill. Each such skill-type is captured by a particular λ payoff-responsiveness value in the logistic model. This converts the game into a game of incomplete information with observable actions. The distribution over the skill-types of each player is common knowledge, but the realized skill-type of each player is private information. Each type of the last mover chooses his last-stage mixed strategy according to his skill-type. The penultimate mover takes the last stage calculation into account and assumes that the last-mover's skill distribution is as given by the commonly known prior. This backward induction procedure can be applied across all stages of the game, where at each stage, each player-type moving there believes that the opponent's skill-type distribution follows the prior. It is straightforward to show that the HAQRE exists and it is unique.

We incorporate a naive disequilibrium version of belief-updating (BU) to define the HAQRE-BU. Naivete is incorporated through the assumption that each player-type assumes she is the only one belief updating. In particular, each player-type assumes that others are playing the unique non-BU HAQRE. In truth, every player-type uses the non-BU HAQRE to update beliefs and plays sequentially rational best responses to the non-BU HAQRE strategies of others. Thus, each player-type can be incorrect in her updated beliefs about each opponent's type distribution as well as her beliefs about their strategies. The benefit of this naivete and disequilibrium assumption is that we do not have to worry about boundedly rational player-types having to account for reputation effects (e.g., as in Kreps et al. (1982)), that may require incongruous sophistication.

For perfect-information games, both HAQRE versions (BU and non-BU) can be solved by simple backward-induction-like recursion without requiring complicated fixed-point calculations of a set of quantal response functions, as needed in QRE, AQRE, and HQRE. Further motivation for the naive-belief-updating assumption comes from the indirect empirical evidence from Kubler and Weizsacker (2004) (KW for short). KW study an information cascades game (a tweaked version of the Anderson and Holt (1997) game); they find that players "behave as if disregarding the fact that their predecessors often use the information that is conveyed by third subjects' decisions" (KW, page 437). Analogously, in HAQRE-BU, we assume each player updates beliefs about opponents' types, but doesn't expect others to do the same.

It is simple to incorporate subjective prior beliefs like truncated downward-looking beliefs in HAQRE and HAQRE-BU. Thus, we define the HAQRE, HAQRE-BU, and their subjective-prior-belief versions. It is straightforward to show that each of these concepts exists, and is unique, both for the logistic versions of QRE as well as for general error distributions.

We extend the HAQRE (with and without BU models) to finitely repeated simultaneous-move games, an important class of extensive-form games. While existence holds, uniqueness doesn't hold here, and the computation of each HAQRE can be complicated since we have to deal with fixed-point calculations. But given the importance of finitely repeated games, our model provides an alternative behavioral theory that can be tested in future research.

We evaluate the HAQRE and related models on experimental data from tests of perfect-information sequential games. In the games we analyze, players with different degrees of experience interacted but players were uncertain about the experience-level of the opponent (Rampal (2020)). The setting provides both an illustration and motivation for our models since heterogeneity in experience-level, a common real-world phenomenon, may well cause differences in skill. We find that accounting for heterogeneity, by comparing HAQRE with AQRE, yields a significant improvement in data-fit. Additionally accounting for belief updating, by comparing HAQRE-BU with HAQRE, also yields a significant improvement in data-fit. Finally, we also find that allowing for subjective prior beliefs can further improve data fit.

2 Related Literature

The model of Noisy Introspection adapted to extensive-form games in Kubler and Weizsacker (2004), which builds on Anderson and Holt (1997), nests the AQRE model (Mckelvey and Palfrey (1998)) as a special case. In the KW model, each player's responsiveness to her own payoff (λ_1) is different from her belief about opponents' payoff responsiveness (λ_2), and her belief about each opponent's beliefs about other opponents' payoff responsiveness (λ_3), and so on.[2] KW find evidence of this heterogeneity in payoff-responsiveness across higher order beliefs. The key difference in the KW model and the present paper is that KW model uncertainty and belief updating about the underlying payoff-relevant state while entertaining no uncertainty about opponents' payoff responsiveness. In contrast, we study uncertainty and belief updating regarding opponents' payoff responsiveness, with the payoff structure known. Other applications of the noisy introspection model have been in simultaneous move games (e.g., Goeree, Louis, and Zhang (2018) study the 11-20 game) unlike the sequential-move games studied here.

Our model of HAQRE for finitely repeated games is related to other studies on the dynamics of beliefs in the bounded rationality literature that deal with learning across *repetitions* of simultaneous-move games (Nyarko and Schotter (2002); Mengel (2014)) or sequential-move games (Ho and Su (2013); Chen, Su, and Zhao (2012); Breitmoser, Tan, and Zizzo (2014)). But these studies often model no heterogeneity (e.g., Breitmoser, Tan, and Zizzo (2014)). Furthermore, we model belief updating about the opponent's skill level, as captured by their payoff responsiveness, which is not studied by these papers. Finally, in our HAQRE model for perfect information games, we allow for updating of beliefs about the opponent's type *within the play* of a sequential-move game, on the basis of preceding moves. By contrast, these papers (and our finitely repeated games model) study belief changes across repetitions of the same game.

Note that we are modeling the beliefs regarding opponents' payoff responsiveness not the payoff themselves, which are certain and common knowledge. Belief updating about the opponent's payoff structure (not payoff responsiveness) is a standard feature of modern game theory of extensive-form games with uncertainty, and it has been studied empirically (e.g., Carrillo and Palfrey (2009)).

In other related work, Friedman (2020) endogenizes the choice of λ or skill of each player in a two-stage model. But unlike our paper, Friedman (2020) does not model uncertainty about payoff responsiveness or updating about the same across the different stages. Friedman (2022) complements the QRE approach, of explaining choice data using noisy actions, by introducing the noisy belief equilibrium (NBE) which allows for noisy beliefs about the opponent's actions but no noise in best response. These noisy beliefs are restricted to satisfy certain axioms. In our HAQRE-BU model, due to the naivete assumption, beliefs about the opponent's strategy can be incorrect. Thus, our motivation for noisy beliefs is quite different, with ours being a disequilibrium approach, which means beliefs can be biased, a feature ruled out by NBE.

[2] also see Goeree and Holt (2004) for an early Noisy Introspection model, and Goeree et al. (2016) for a discussion of this literature.

3 Model

Consider an arbitrary finite sequential-move game with perfect information and perfect recall Γ, which is known henceforth as the base game. The base game Γ is defined as a collection of the following components.[3] A finite player-set, N. A finite set of finite histories, H. A history $h \in H$ is a finite sequence of (pure) actions, e.g., $h = (a^s)_{s=1,\ldots,S}$. The s^{th} action, a^s, is said to be taken at the s^{th} stage of the game. Subsequences/subhistories are defined in the usual fashion. Z denotes the set of terminal histories, with the elements of Z denoted as z. Each sequence/history in $(H \backslash Z)$ is a singleton information set. A set of possible pure actions in the game, A, and an action correspondence $A(.)$ which maps $h \in (H \backslash Z)$ to an action set at h, $A(h) \equiv \{a \in A \mid (h, a) \in H\}$. A player function $P(.)$ mapping non-terminal histories to the player moving there. Finally, a Bernoulli utility function u_i for each player $i \in N$; u_i maps terminal histories to real numbers. Thus, Γ is defined as $\{N,\, H,\, P,\, A,\, \{u_i\}_{i \in N}\}$.

The Agent Quantal Response Equilibrium (AQRE) model of Mckelvey and Palfrey (1998) extends the Quantal Response Equilibrium of McKelvey and Palfrey (1995) to extensive form games. The AQRE model introduces a *payoff disturbance* error term to the expected payoff associated with each action of each player at each possible move of that player. Let H^i be all the information sets where player i moves, let the k^{th} information set in H^i be denoted h^i_k. Let $\overline{u}_i(a, b \mid h^i_k)$ be the expected payoff of player i from playing action a at h^i_k given that the behavioral strategies for all players moving at subsequent information sets is as per the behavioral strategy profile b. Then, player i perceives that her payoff after action a is: $\hat{u}_i(a, b \mid h^i_k) = \overline{u}_i(a, b \mid h^i_k) + \epsilon_{ika}$, where ϵ_{ika} is the value of the error term after action a from history h^i_k.

In an AQRE b, each player i, at each of her moves h^i_k takes the action that she perceives will yield the highest possible expected payoff among the actions available at that move, given that all players are playing according to the behavioral strategy profile b. Thus, in an AQRE b, the probability that i will play a at h^i_k is given by

$$b(a|h^i_k) = Pr(a \text{ yields max expected payoff to } i \text{ among } A(h^i_k)|b, h^i_k) =$$

$$= Pr[(\epsilon_{ika})_{a \in A(h^i_k)}|\overline{u}_i(a, b \mid h^i_k) + \epsilon_{ika} \geq \overline{u}_i(a', b \mid h^i_k) + \epsilon_{ika'} \ \forall a' \in A(h^i_k)]. \quad (1)$$

In the common logit formulation of AQRE, ϵ_{ika} is assumed i.i.d., for all i, k, and a, according to type I extreme value distribution with c.d.f. $F(\epsilon_{ika}) = e^{-e^{-\lambda \epsilon_{ika}}}$. Thus, the probability

[3] See Osborne and Rubinstein (1994) for a more complete definition of certain terms.

$b(a|h)$ assigned to action a at any history $h \in H^i$ in a logit-AQRE is:

$$b(a|h) = Pr(a \text{ yields max expected payoff to } i \text{ among } A(h)|b,h) =$$

$$= \frac{exp(\lambda \bar{u}_i(a \mid b,h))}{\sum_{a' \in A(h)} exp(\lambda \bar{u}_i(a' \mid b,h))}. \tag{2}$$

Informally, λ captures a player's payoff responsiveness or skill: for a given $\lambda \geq 0$, each action receives a probability weight proportional to the expected payoff $\bar{u}_i(a, b \mid h)$ from that action relative to other actions in $A(h)$. As λ increases, each player at each move puts more probability weight on the response whose expected payoff is the highest. So, the parameter λ generates a unique logit-AQRE b.

In this paper we model the scenario where there is uncertainty about the skill/payoff-responsiveness or, more technically, error distributions of other players. This means that the base game of complete and perfect information is converted into a game of incomplete information with observable actions Γ^I. In Γ^I, each player of the base game can be one of several types where type implies a particular error distribution (henceforth referred to as "skill") with which she perceives her expected payoff. Note that Γ^I can also be reinterpreted as a game of imperfect information with uncertainty about Nature's actions in selecting the skill-type of each player. In Γ^I each player knows her own type, but each player is uncertain about her opponent's type. Since the construction of an incomplete-information game is standard, we provide a definition that skips over several formalities.

First consider the logistic case. The incomplete information game Γ^I can be defined from the base game Γ as follows. For each player $i \in N$ of Γ, there is a set of player-types $\Lambda_i = \{\lambda_i^1, ..., \lambda_i^{m(i)}\}$, where 1,.., $m(i)$ are natural numbers, $m(i)$ denotes the total number of types of player i, and for natural numbers n_1 and n_2, $n_1 < n_2 \leq m(i)$ implies $\lambda_i^{n_1} < \lambda_i^{n_2}$. The inequality $\lambda_i^{n_1} < \lambda_i^{n_2}$ can be interpreted as: type $\lambda_i^{n_1}$ is less skillful than $\lambda_i^{n_2}$. The common knowledge prior probability distribution over the types of each player i is given by π_i for $i \in N$, and for $i \neq j$, π_i is independent of π_j for all players $i, j \in N$.

3.1 Example 1: Logistic HAQRE

Consider an example of a two-player base game Γ illustrated in Figure 1. Suppose instead of a common λ for both players, as in the logit-AQRE, we have two possible levels of λ for either player. For player 1, $\lambda_1 \in \Lambda_1 = \{\lambda_1^L, \lambda_1^H\}$, and for player 2, $\lambda_2 \in \Lambda_2 = \{\lambda_2^l, \lambda_2^h\}$. Further, suppose the probability distributions on Λ_1 and Λ_2, given by π_1 and π_2, are common knowledge. We now describe the logit version of heterogeneous AQRE, or HAQRE for short, with two possible λ-values for each player; this generates the imperfect information game

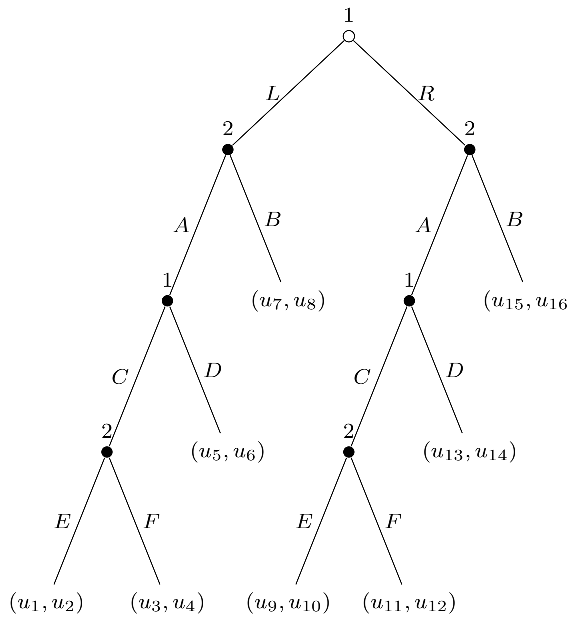

Figure 1: A base game example

Γ^I (not shown) corresponding to the base game Γ in Figure 1.

Step 1: We will construct the HAQRE recursively, backwards from the last stage to the first stage. Player 2 moves at the fourth (last) stage. So, consider $h4$, an arbitrary non-terminal history reaching the fourth stage (with three preceding actions). Recall that there are two types of player 2: $\lambda_2 \in \{\lambda_2^l, \lambda_2^h\}$. The probability that any action $a \in A(h4)$ will be played at $h4$, is determined by (2) and the appropriate λ_2 value. So, if $\lambda_2 = \lambda_2^h$, then player-2 plays $a \in A(h4)$ with the probability

$$b(a|h4, \lambda_2^h) = \frac{exp(\lambda_2^h \bar{u}_2(a \mid h4))}{\sum_{a' \in A(h4)} exp(\lambda_2^h \bar{u}_2(a' \mid h4))} \tag{3}$$

and if $\lambda_2 = \lambda_2^l$ then

$$b(a|h4, \lambda_2^l) = \frac{exp(\lambda_2^l \bar{u}_2(a \mid h4))}{\sum_{a' \in A(h4)} exp(\lambda_2^l \bar{u}_2(a' \mid h4))}. \tag{4}$$

Notice that at the last stage, we didn't need any input about the rest of the strategy profile to calculate payoffs from any action; thus, $\bar{u}_2(a \mid h4)$ is the terminal payoff in Γ from choosing a at $h4$. Clearly, this holds true for all actions at "penultimate nodes", i.e., actions which yield a terminal history. Thus, in Figure 1, (3) and (4) can be used to calculate the probabilities with which the λ_2^h and λ_2^l types of player 2 will play E and F at stage-4.

Step 2: Consider player 1 moving at an arbitrary move $h3$ at stage-3. In the spirit of backward induction, we will assume player 1 knows that player 2 will play according to $b(a|h4, \lambda_2)$ in stage 4, for $\lambda_2 \in \{\lambda_2^h, \lambda_2^l\}$, as calculated in (3) and (4), for all a in $A(h4)$ and for all $h4$ in stage 4. Let b^4 denote this strategy profile, constructed in step 1, for either type of player-2 at stage 4. Note that player 1 doesn't know the type of player 2, and we assume player 1 believes that the probabilities with which player 2 is λ_2^h or λ_2^l are given by the prior, $\pi_2(\lambda_2^h)$ and $\pi_2(\lambda_2^l)$, respectively. Thus, for an action a at $h3$ that doesn't terminate the history, player 1 uses the prior π_2 and b^4 to calculate the expected payoff of a; this expected payoff can be denoted as $\bar{u}_1(a \mid h3, b^4, \pi_2)$. Thus, using (2), the probabilities with which the types λ_1^H and λ_1^L of player 1 will choose the different actions available at $h3$ are given by:

$$b(a \mid h3, \lambda_1) = \frac{exp(\lambda_1 \bar{u}_1(a \mid h3, b^4, \pi_2))}{\sum_{a' \in A(h3)} exp(\lambda_1 \bar{u}_1(a' \mid h3, b^4, \pi_2))} \text{ for } \lambda_1 \in \{\lambda_1^H, \lambda_1^L\}. \tag{5}$$

For actions that terminate a history (e.g., action D) the expected payoff is simply determined by the action and the history, like in step 1.

Step 3: Let the strategy profile constructed in steps 1 and 2 for stages 4 and 3 be denoted b^3. Consider player 2 moving at an arbitrary move $h2$ at stage-2; her expected payoff from an action depends on the subsequent strategy profile b^3 and her beliefs over player-1's type (given by the prior, π_1): $\bar{u}_2(a \mid h2, b^3, \pi_1)$. Notice that player-2 of type λ_2 knows her own type, thus she knows the probabilities with which she will play stage-4 actions, as specified by the components of b^3 pertaining to her own value of λ_2. Regarding player-1, player-2 needs to know the components of b^3 specifying the strategies of both player-1 types at stage-3, as well as the prior probability distribution over player-1's types. Thus, analogous to (5), for $\lambda_2 \in \{\lambda_2^h, \lambda_2^l\}$, we have:

$$b(a \mid h2, \lambda_2) = \frac{exp(\lambda_2 \bar{u}_2(a \mid h2, b^3, \pi_1))}{\sum_{a' \in A(h2)} exp(\lambda_2 \bar{u}_2(a' \mid h2, b^3, \pi_1)} \text{ for } \lambda_2 \in \{\lambda_2^h, \lambda_2^l\}. \tag{6}$$

Step 4: Consider player 1 moving at $h1$ at stage-1. The expected payoff of player-1 of type λ_1 from an action a at $h1$ can be denoted as $\bar{u}_1(a \mid h1, b^2, \pi_2)$, where b^2 denotes the strategy profile for stages 2-4, calculated in steps 1-3. Note that player 1 calculates the second-stage strategy component of b^2 as if she was in player-2's mind at player-2's stage-2 move. Therefore, at stage-1, player 1 calculates b^2 exactly as described in step 2 (which includes accounting for π_1). Similar to preceding steps, we have:

$$b(a \mid h1, \lambda_1) = \frac{exp(\lambda_1 \bar{u}_1(a \mid h1, b^2, \pi_2))}{\sum_{a' \in A(h1)} exp(\lambda_1 \bar{u}_1(a' \mid h1, b^2, \pi_2))} \text{for } \lambda_1 \in \{\lambda_1^H, \lambda_1^L\}. \tag{7}$$

The HAQRE is the strategy profile b constructed in the Steps 1-4 above from the behavioral strategy for each stage of the game.■

There are two notable issues in Example 1. First, note that we have not modeled possible belief updating that players can do about the opponent's λ on the basis of the opponent's preceding moves. This is an issue unique and important for sequential-move games. Furthermore, unlike the AQRE where players at different moves are treated as essentially different players, in the HAQRE the same player at a different move is recognized as the same player. Thus, exploring belief updating is both possible and important. We do this by defining the HAQRE-BU (with belief updating) formulation in the next section. The second issue is that, as in Rogers, Palfrey, and Camerer (2009) formulation of truncated heterogeneous QRE, it is possible that some player-types with a certain λ may only allow for the possibility of opponents with λ lower than a certain threshold, i.e., a truncated HAQRE. We define and discuss the truncated HAQRE in the next section.

4 Heterogeneous AQRE and belief updating

Consider an arbitrary S-staged finite base game of complete and perfect information Γ, and the associated incomplete-information game with observable actions: Γ^I. In this section, we define the HAQRE, HAQRE-BU, and the truncated HAQRE for Γ^I with the logistic error distribution and "general" error distributions for finite perfect information base games.

4.1 Logistic Heterogeneous AQRE without belief updating

The logistic version of HAQRE (without belief updating) is a behavioral strategy profile for Γ^I which specifies for each type of each player, the (possibly mixed) action they will take at each move. The definition is a generalization of the HAQRE construction in Example 1.

Definition 1 (Logistic HAQRE). *The Logistic HAQRE of an S-staged Γ^I is a behavioral strategy profile b which is defined by recursive construction as follows.*
Step 1. *Consider an arbitrary stage-S history hS with $(S-1)$ preceding actions. The player moving there is $P(hS)$. The probability that each type $\lambda_{P(hS)}$ of $P(hS)$ chooses an arbitrary action $a \in A(hS)$ is given by*

$$b(a|hS, \lambda_{P(hS)}) = \frac{exp[\lambda_{P(hS)}\bar{u}_{P(hS)}(a \mid hS)]}{\sum_{a' \in A(hS)} exp[\lambda_{P(hS)}\bar{u}_{P(hS)}(a' \mid hS)]}. \tag{8}$$

Step s, for $s \in \{2, .., S\}$***.*** *Consider an arbitrary stage-$(S-s+1)$ history $h(S-s+1)$; call it h_s for short. The player moving there is $P(h_s)$. At step s, we have completed steps*

1 through $(s-1)$*, in which we have calculated* $b(a|hs', \lambda_{P(hs')})$ *for any non-terminal history* hs' *at stage-*s' *with* $s' \in \{S-s+2,..,S\}$ *for all types* $\lambda_{P(hs')}$ *of* $P(hs')$ *for all* $a \in A(hs')$*. Let these behavioral strategies constructed in steps 1 through* $(s-1)$ *be denoted as* b^s*. Then, type* $\lambda_{P(h_s)}$ *of* $P(h_s)$ *can calculate the expected payoff of her action* $a \in A(h_s)$ *using* b^s*, the probability distributions over other players' types,* $\pi_{-P(h_s)}$*, and her own type that she knows; let this expected payoff be denoted:*

$$\bar{u}_{P(h_s)}(a \mid h_s, b^s, \pi_{-P(h_s)}, \lambda_{P(h_s)}).$$

So, in the Logistic HAQRE, the probability that each type $\lambda_{P(h_s)}$ *of* $P(h_s)$ *chooses an arbitrary action* $a \in A(h_s)$ *is given by*

$$b(a|h_s, \lambda_{P(h_s)}) = \frac{exp[\lambda_{P(h_s)}\bar{u}_{P(h_s)}(a \mid h_s, b^s, \pi_{-P(h_s)}, \lambda_{P(h_s)})]}{\sum_{a' \in A(h_s)} exp[\lambda_{P(h_s)}\bar{u}_{P(h_s)}(a' \mid h_s, b^s, \pi_{-P(h_s)}, \lambda_{P(h_s)})]} \blacksquare \qquad (9)$$

The definition yields certain properties of the Logistic HAQRE.

Proposition 1 (existence and uniqueness). *For any finite game of incomplete information with observable action generated from a base game of complete and perfect information, there exists a unique logistic HAQRE.*

Proof. Consider Definition 1 specifying the recursive construction of the Logistic HAQRE b for an arbitrary Γ^I constructed from an underlying base game Γ. In step 1, for each type $\lambda_{P(hS)}$ of $P(hS)$ and for any action $a \in A(hS)$, $b(a|hS, \lambda_{P(hS)})$ is well defined and uniquely determined by (8). Furthermore, in step-s, with $s \in \{2,..,S\}$ we consider an arbitrary stage-$(S-s+1)$ history $h(S-s+1)$, or h_s for short, and for each type $\lambda_{P(h_s)}$ of $P(h_s)$ and for each action $a \in A(h_s)$ we define $b(a|h_s, \lambda_{P(h_s)})$ using (9). Note that the probability with which a is played at h_s by $\lambda_{P(h_s)}$, $b(a|h_s, \lambda_{P(h_s)})$, is well defined and uniquely determined by (9). Thus, repeating step s from step 2 until step S yields a well defined and unique logistic HAQRE for Γ^I. $\blacksquare$

It is straightforward to define the logistic HAQRE for general error distributions (while still maintaining the i.i.d. error assumption) and to show that such a "general HAQRE" exists. This point is summarized in Fact 1. We refer the reader to the Appendix to see this model.

Fact 1: For every Γ^I, the general HAQRE exists, and it is unique.

4.2 Logistic HAQR (dis)Equilibrium with belief updating

In a sequential-move game with incomplete information, as our Γ^I, if there is an opportunity to observe preceding actions, then there is an opportunity to form updated beliefs about the opponent's type. In the logistic HAQRE with belief updating (HAQRE-BU) defined in this subsection, we incorporate this belief updating.

Inspired by the findings of Kubler and Weizsacker (2004) (discussed in the Introduction) the belief updating we model in the HAQRE-BU is of a naive nature, which yields a disequilibrium concept. We assume that each player-type updates beliefs after observing preceding actions, but each player-type assumes she is the only one updating beliefs. In effect, she does not account for opponents' types performing the analogous belief updating. Thus, player-types don't account for the "reputation effects" of their actions (see Kreps et al. (1982) and the subsequent literature on reputation). Instead, each player-type assumes that each type of each opponent holds prior beliefs at all of her/his moves.

The rationale for this choice is incorporating simplicity in the reasoning process of each player-type. In particular, it seems that a full scale belief updating and reputation model would involve a very high level of cognitive sophistication from player-types, a feature that seems incompatible with bounded-rational players who perceive their expected payoffs with mistakes.

Each player-type needs a reference strategy profile (that she thinks opponent-types are playing) to update beliefs about which type of each opponent could have played the actions she observes before her move. We assume that each player-type assumes that all other player-types are always following the HAQRE strategy profile since they are playing using prior beliefs throughout the game. Note that this is a generically incorrect assumption by player-types that everyone else follows the HAQRE strategy profile while each one actually plays the HAQRE-BU strategy profile. Thus, it should be noted that HAQRE-BU is really a disequilibrium approach: player-types are generically wrong about others' beliefs and their strategies.

For ease-of-use, we first state the definition of HAQRE-BU informally. The HAQRE-BU of Γ^I is an assessment with two components: (i) a profile of beliefs, and (ii) a strategy profile. The HAQRE-BU satisfies the following.

(1) The belief profile is such that, at each information set, the beliefs are calculated using Bayes' rule with the following inputs: the common-knowledge prior, the unique HAQRE of Γ^I, and the actions preceding that information set.

(2) The strategy profile is such that, at each information set, the mixture over the action set chosen is a logistic response to the expected payoffs from the actions in that set. These expected payoffs are calculated using the HAQRE-BU beliefs (calculated in (1)), assuming that other player-types are playing the unique HAQRE of Γ^I, and that the player-type

herself will follow her HAQRE-BU strategies at her subsequent moves.[4] Due to the latter, a backward-induction recursive procedure is needed for the construction of HAQRE-BU.

Now we formally define this logistic HAQRE-BU with naive beliefs.

Definition 2 (Logistic HAQRE-BU). *Logistic HAQRE with belief updating (HAQRE-BU) is a behavioral strategy profile b and a profile of beliefs $\mu = (\mu_{\lambda_i})_{\lambda_i \in \Lambda_i \text{ and } i \in N}$ for Γ^I. We first define μ. Consider the unique Logistic HAQRE strategy profile of Γ^I, denoted b^H. Throughout Definitions 2 and 2a fix an arbitrary information set hs with some player-type $\lambda_{P(hs)}$ moving there. The HAQRE-BU belief of $\lambda_{P(hs)}$ is calculated through Bayesian updating using the prior distribution and the HAQRE strategy profile b^H. Thus, we have for all $i \neq P(hs)$ and all $\lambda_i \in \Lambda_i$,*

$$\mu_{\lambda_{P(hs)}}(\lambda_i|hs, b^H, \pi) = \frac{Pr(hs \text{ and } \lambda_i|\pi, b^H)}{\sum_{\lambda'_i \in \Lambda_i} Pr(hs \text{ and } \lambda'_i|\pi, b^H)}. \tag{10}$$

Now we construct the HAQRE-BU strategy profile b. Conditional on reaching the information set hs, $\lambda_{P(hs)}$'s (referred to as "her") expectation of payoffs from an arbitrary action $a \in A(hs)$ is a function of: (i) her beliefs over opponents' types, which are given by $\mu_{\lambda_{P(hs)}}$; (ii) her expectation about all opponent-types' behavioral strategies at subsequent moves, which are as per the HAQRE b^H; and finally (iii) her expectation about her own actions at her own subsequent moves (if such moves exist), which we denote as $b^{\lambda_{P(hs)}}$ (Definition 2a, below). Let $\lambda_{P(hs)}$'s expectation of payoffs from an arbitrary action $a \in A(hs)$ be denoted

$$\bar{u}_{P(hs)}(a \mid hs, b^H, \mu_{\lambda_{P(hs)}}, b^{\lambda_{P(hs)}}).$$

Then in HAQRE-BU, $\lambda_{P(hs)}$ at her move hs chooses $a \in A(hs)$ with probability

$$b(a|hs, \lambda_{P(hs)}) = \frac{exp[\lambda_{P(hs)}\bar{u}_{P(hs))}(a \mid hs, b^H, \mu_{\lambda_{P(hs)}}, b^{\lambda_{P(hs)}})]}{\sum_{a' \in A(hs)} exp[\lambda_{P(hs)}\bar{u}_{P(hs))}(a' \mid hs, b^H, \mu_{\lambda_{P(hs)}}, b^{\lambda_{P(hs)}})]}. \tag{11}$$

Finally, we define $b^{\lambda_{P(hs)}}$, the expectations of $\lambda_{P(hs)}$ about her own actions at her own subsequent moves. This is constructed recursively: start from the last moves of $\lambda_{P(hs)}$, use the logit rule to specify probabilities of actions there (with the expected payoffs calculated exactly as done in Definition 2 using b^H and $\mu_{\lambda_{P(hs)}}$), fix that, and work backwards to preceding moves of $\lambda_{P(hs)}$, and continue this recursion all the way to hs. Due to finiteness of the game, this procedure is well defined. We state this formally in Definition 2a.

[4] When performing the backward-induction calculations, each player-type accounts for the fact that, at subsequent stages, her own beliefs will have been further updated and will be possibly different from her current beliefs.

Definition 2a (defining $b^{\lambda_{P(hs)}}$): *First, consider those moves hs^L of $\lambda_{P(hs)}$ such that there are no subsequent moves of $\lambda_{P(hs)}$ in any possible subsequent histories from hs^L.*[5] *For all $a \in A(hs^L)$, fix*

$$b^{\lambda_{P(hs)}}(a|hs^L, \lambda_{P(hs)}) = \frac{exp[\lambda_{P(hs)}\bar{u}_{P(hs))}(a \mid hs^L, b^H, \mu_{\lambda_{P(hs)}})]}{\sum_{a' \in A(hs^L)} exp[\lambda_{P(hs)}\bar{u}_{P(hs))}(a' \mid hs^L, b^H, \mu_{\lambda_{P(hs)}})]}.$$

[6]*Next, for each hs^L, consider the move of $\lambda_{P(hs)}$ immediately preceding hs^L on the path from hs to hs^L; label this hs^{L-1}.*[7] *For all $a \in A(hs^{L-1})$:*

$$b^{\lambda_{P(hs)}}(a|hs^{L-1}, \lambda_{P(hs)}, b^{\lambda_{P(hs)}}(hs^L)) =$$

$$= \frac{exp[\lambda_{P(hs)}\bar{u}_{P(hs))}(a \mid hs^{L-1}, b^H, \mu_{\lambda_{P(hs)}}, b^{\lambda_{P(hs)}}(hs^L))]}{\sum_{a' \in A(hs^{L-1})} exp[\lambda_{P(hs)}\bar{u}_{P(hs))}(a' \mid hs^{L-1}, b^H, \mu_{\lambda_{P(hs)}}, b^{\lambda_{P(hs)}}(hs^L))]} \qquad (12)$$

where we use $b^{\lambda_{P(hs)}}(hs^L)$ constructed in the first step for all hs^L. Repeating the second step for the move of $\lambda_{P(hs)}$ immediately preceding each hs^{L-1}, and thereafter proceeding similarly and recursively to hs yields $b^{\lambda_{P(hs)}}$.

Proposition 2 (existence and uniqueness of logistic HAQRE-BU). *For any finite game of incomplete information with observable action generated from a base game of complete and perfect information, there exists a unique logistic HAQRE-BU.*

Proof. First, note that the HAQRE b^H used to construct the logistic HAQRE-BU (b, μ) exists, and is unique. Thus, by (10), μ is unique and well defined since each node of Γ^I is reached with strictly positive probability under b^H. Next, note that if $b^{\lambda_{P(hs)}}$ is well defined and unique for each hs and $\lambda_{P(hs)}$, then the proof is complete.

To see that $b^{\lambda_{P(hs)}}$ is indeed well defined and unique, note the following. The first step of the construction of $b^{\lambda_{P(hs)}}$ utilizes the well-defined and unique b^H and $\mu_{\lambda_{P(hs)}}$; this yields a well defined and unique $b^{\lambda_{P(hs)}}(a|hs^L)$ for all $a \in A(hs^L)$. Consequently, in the recursive construction of $b^{\lambda_{P(hs)}}$, at each subsequent step, the required components of $b^{\lambda_{P(hs)}}$ are well defined and unique. Thus, the recursive construction of $b^{\lambda_{P(hs)}}$ yields a well defined and unique $b^{\lambda_{P(hs)}}$, which means that the logistic HAQRE-BU for Γ^I, (b, μ), is well defined (i.e.,

[5]*Formally, hs^L is that history such that hs is a sub-history of hs^L and there is no history $hs' \neq hs^L$ of Γ^I such that $\lambda_{P(hs)}$ moves at hs' and hs^L is a strict sub-history of hs'.*

[6]*Note that $\mu_{\lambda_{P(hs)}}(hs^L)$ implies that the component of the beliefs of $\lambda_{P(hs)}$ used here is the HAQRE-BU belief distribution of $\lambda_{P(hs)}$ at hs^L. That is, at hs, $\lambda_{P(hs)}$ recognizes how her beliefs would have been updated by the time she moves at hs^L.*

[7]*Formally, hs^{L-1} is that history such that (i) hs is a sub-history of hs^{L-1}, (ii) hs^{L-1} is a subhistory of hs^L, and there is no sub-history $hs' \neq hs^{L-1}$ of hs^L such that $\lambda_{P(hs)}$ moves at hs', hs is a sub-history of hs', and hs^{L-1} is a sub-history of hs'.*

it exists) and (b, μ) is unique. ■

It is straightforward to define the logistic HAQRE-BU for general error distributions (while still maintaining the i.i.d. error assumption) and to show that such a "general HAQRE-BU" exists. This point is summarized in Fact 2. We refer the reader to the Appendix to see this model.

Fact 2: For every Γ^I, the general HAQRE-BU exists, and it is unique.

4.3 Subjective and Truncated logit-HAQRE

It may be the case that each type of each player has subjective prior beliefs over each opponent's types. We denote this as subjective HAQRE, or SHAQRE. A special case of interest is where each player-type believes that each opponent-type is less skillful than herself. This is of special interest since it is the analogous extensive-form version of the Truncated Heterogeneous QRE (THQRE), which Rogers, Palfrey, and Camerer (2009) define. While Rogers, Palfrey, and Camerer (2009) define the THQRE for arbitrary truncation points, we focus on only "downward looking" beliefs in our analogous model. They show that THQRE (with downward looking beliefs) helps connect the QRE and Level-k models. We define our analogous Truncated HAQRE below, and put the more general Subjective HAQRE (which can be used to model arbitrary truncation points), with and without belief updating, in the Appendix.

To repeat the motivation for the truncated model provided by Rogers, Palfrey, and Camerer (2009), there are some reasons why truncated beliefs represent a reasonable constraint on players' beliefs: one is that if a player with low λ can compute what players with higher values of λ do, she may well target such higher-type behavior for herself. Second, evidence from the psychology literature indicates that people are often overconfident (e.g. Kahneman and Tversky (1973); Camerer and Lovallo (1999)). Furthermore, Kubler and Weizsacker (2004) find that subjects on average attribute a lower response precision to their opponents than they have themselves. Finally, the benefits of considering the existence of more λ-types when computing expected payoffs might not justify its cognitive costs.

The Truncated Logistic-HAQRE is designed to capture the idea that player i of skill-type or payoff-responsiveness λ_i^n assumes that each player j has a lower type, i.e. $\lambda_j \leq \lambda_i^n$ holds for all $\lambda_j \in \Lambda_j$, and all j. Let player-type λ_i^n's beliefs be denoted by $\mu_{\lambda_i^n}$. If some player j has some types that are lower than λ_i^n, then λ_i^n believes that the probability of each such type is equal to that type's probability conditional on being a type lower than λ_i^n (where these calculations use the true prior); this is specified by (13) below. In case the true distribution of types of a given opponent does not contain any type lower than

λ_i^n, then λ_i^n considers that the opponent's type is the same as her own type. Assumption 1 specifies these beliefs of each player-type in the Truncated Logistic-HAQRE.

Assumption 1. In the Truncated Logistic-HAQRE, type λ_i^n of player i believes the following probability distribution on each player-j's types. If there is no $\lambda_j \leq \lambda_i^n$ in Λ_j, then λ_i^n's belief about j's type is given by $\mu_{\lambda_i^n}(\lambda_j^m) = 1$ if and only if $\lambda_j^m = \lambda_i^n$. Otherwise, λ_i^n's belief about j's type is as follows:

$$\mu_{\lambda_i^n}(\lambda_j^m) = \frac{\pi_j(\lambda_j^m)}{\sum_{\lambda_j^{m'} \leq \lambda_i^n} \pi_j(\lambda_j^{m'})} \text{ if } \lambda_j^m \leq \lambda_i^n; \text{ otherwise } \mu_{\lambda_i^n}(\lambda_j^m) = 0. \tag{13}$$

Note that we are again in disequilibrium territory since beliefs of player-types can be incorrect. Furthermore, each player-type's beliefs about the distribution of types across players are distinct from the beliefs of each other player-type, except for those opponent-types who share her own payoff responsiveness. Since we are in a setting where players are boundedly rational (specifically, players perceive their payoffs with mistakes), we want to avoid complications related to higher-order beliefs. To that end, we assume the following. In this Truncated Logistic-HAQRE (THAQRE for short) setting, we assume that when any given player-type λ_i^n calculates her expected payoff from a certain action at a certain move of hers, she assumes that all player-types moving at subsequent moves will act "as if" they hold the same beliefs over other players' types as her own beliefs about those players' types.

Note that higher-order beliefs are not an explicit part of the definition of a THAQRE, yet these are tacitly present while applying THAQRE to data. This inaccuracy of tacit higher-order beliefs, on top of inaccurate first-order beliefs, brings us into further disequilibrium territory. Now, we formally define the Truncated Logistic-HAQRE using the Logistic HAQRE.

Definition 3. *The Truncated Logistic-HAQRE is a behavioral strategy profile b^T for Γ^I. For each player-type λ_i^n, let $\mu_{\lambda_i^n}$ be the subjective prior of λ_i^n satisfying Assumption 1. Let $\Gamma^{I\prime}$ be a game identical to Γ^I except that the common-prior of Γ^I is replaced by $\mu_{\lambda_i^n}$ instead of π. The Truncated Logistic-HAQRE behavioral strategy of λ_i^n for Γ^I, denoted $b^T(\lambda_i^n)$, is given by λ_i^n's Logistic HAQRE behavioral strategy $b'(\lambda_i^n)$ for $\Gamma^{I\prime}$.*

Notice that according to Definition 3, to construct the Truncated Logistic-HAQRE, one can follow the steps $s \in \{1, .., S\}$ of the construction of the logistic HAQRE in Definition 1, except the following change. Instead of the true probability distributions over other players' types, $\pi_{-P(h_s)}$, being used by $\lambda_{P(h_s)}$ to calculate his expected payoff from each action $a \in A(h_s)$, Assumption 1 is used to define $\mu_{\lambda_{P(h_s)}}$, the probability distribution over opponents' types believed by $\lambda_{P(h_s)}$ that he uses to calculate his own expected payoff from

a. But also, $\lambda_{P(h_s)}$ acts as if tacitly using $\mu_{\lambda_{P(h_s)}}$ to form higher-order beliefs. That is, each player-type tacitly uses (in the recursive backward induction procedure) his own subjective beliefs to model how the player in the next stage is modeling the player who follows the latter, and so on. This is why we have to replace the common-prior of Γ^I with each player-type's own subjective belief to define that player-type's strategy in the Truncated Logistic-AQRE (THAQRE). Due to the use of the Logistic HAQRE in defining THAQRE, and due to Proposition 1 (existence and uniqueness of the Logistic HAQRE), we have the following.

Proposition 3 (existence and uniqueness of the Truncated Logistic HAQRE). *For any finite game of incomplete information with observable action generated from a base game of complete and perfect information, there exists a unique Truncated Logistic HAQRE.*
Proof: For each player-type λ_i^n, $b^T(\lambda_i^n)$ is given by λ_i^n's Logistic HAQRE behavioral strategy $b'(\lambda_i^n)$ for $\Gamma^{I\prime}$. Since $b'(\lambda_i^n)$ exists and is unique by Proposition 1, Proposition 3 holds.■

5 HAQRE for repeated games

An important class of extensive-form games is finitely repeated simultaneous-move games. Here we shall focus on finitely repeated games with observable actions, where the actions taken by each player in each stage are common knowledge before all subsequent stages. This class of games includes the finitely repeated prisoner's dilemma, Cournot, hawk-dove, matching-pennies, and many other important finitely repeated games. To reiterate, we are interested in modeling scenarios where such repeated games are being played by players with possibly different skill levels, and play in the preceding stages provides players an opportunity to update beliefs about the skill level of each opponent.

Formally, the n-staged repeated game consists of the stage game G repeated n times. The stage game is defined as: $G \equiv \{N; (A_i)_{i\in N};\ u_i : A \to \mathbb{R}$ for each player $i\}$, which contains the following—the player set N, the action set A_i for each player i, and the payoff function u_i mapping $A = \times_{i\in N} A_i$ to real numbers. A k-staged history is $(a^1, ..., a^{k-1})$, where a^s denotes the action profile played in the s^{th} stage. At each stage, i.e., at each repetition of G, the history of all the preceding stages is common knowledge. Thus, the base game is $\Gamma = \{G, n\}$, which means Γ is the stage game G repeated n times. We will assume that there is no discounting within Γ, and a player's payoff in Γ is the sum of that player's payoff in each of the n stages.[8a]

To model heterogeneity and uncertainty about each opponent's skill, we convert Γ into the incomplete information game Γ^I. The incomplete information game Γ^I can be defined

[8a,b]The general case is straightforward to model, but it is not included in this draft.

from the base game Γ as done earlier, with each player $i \in N$ of Γ replaced by a set of player-types. This set of player-types is given by $\Lambda_i = \{\lambda_i^1, ..., \lambda_i^{m(i)}\}$ in the logistic case.[8b] The common knowledge prior probability distribution over the types of each player i is given by π_i for $i \in N$, and for $i \neq j$, π_i is independent of π_j for all players $i, j \in N$.

5.1 Defining the HAQRE for repeated games

For each stage $G^I(s)$ of Γ^I, the HAQRE (without BU) specifies a discrete version of the HQRE of Rogers, Palfrey, and Camerer (2009). That is, the HAQRE is a behavioral strategy profile b that specifies that each type of player i, say λ_i^r, chooses actions with probabilites proportional to the expected payoff from each action. This expected payoff is calculated on the basis of the prior beliefs over other players' types (given that there is no belief updating), and given their respective HAQRE behavioral strategies in $G^I(s)$. Due to backward induction, without belief updating, the behavioral strategies in any stage are independent of the history of play reaching that stage. Thus, the HAQRE (without belief updating) is simply a discrete HQRE in each stage; no player-type has to consider different "paths" originating from different actions as in the HAQRE corresponding to perfect information games. The formal definition of the HAQRE b for repeated games is as follows.

Definition 4 (HAQRE for finitely repeated games). *For an arbitrary stage of Γ^I, denote by $\bar{u}_{\lambda_i^r}(a_i \mid b_{-\lambda_i^r}, \pi)$ the expected payoff of λ_i^r from action a_i in that stage, given that other player-types' behavioral strategy is given by $b_{-\lambda_i^r}$, and π denotes the prior. For each player $i \in N$, each player-type $\lambda_i^r \in \Lambda_i$ chooses each action $a_i \in A_i$ with the probability*

$$b^{\lambda_i^r}(a_i|b_{-\lambda_i^r}, \pi) = \frac{exp[\lambda_i^r \bar{u}_{\lambda_i^r}(a_i \mid b_{-\lambda_i^r}, \pi)]}{\sum_{a_i' \in A_i} exp[\lambda_i^r \bar{u}_{\lambda_i^r}(a_i' \mid b_{-\lambda_i^r}, \pi)]}. \tag{14}$$

Since Definition 4 is a straightforward application of the HQRE of Rogers, Palfrey, and Camerer (2009), by their Theorem 1, we have the following Proposition.

Proposition 4: *The HAQRE for finitely repeated games exists.*

The HAQRE may not be unique, since even the HQRE for a stage game may not be unique. This can be seen by considering the hawk-dove game, which has two Nash equilibria. If there is little or no heterogeneity in λ there may be multiple HQRE, each in the "neighbourhood" of the multiple Nash equilibria of games like the hawk-dove game.

Next, we define the HAQRE-BU (with belief updating) for repeated games. We again model naive belief updating and take a disequilibrium approach. That is, each player-type incorrectly believes he is the only one performing belief updation. Specifically, each player-type incorrectly believes that all others hold prior beliefs throughout the game. This helps us avoid issues related to reputation building.

Thus, each player-type expects that in each stage, others' strategies are selected from some HQRE (where the HQRE uses the prior). This expected strategy profile is what the player-type uses to perform belief updating using Bayes' rule. Within this setup, in a given stage, a player-type's subjective expected payoff from an action is calculated using her updated beliefs, and her HQRE-based expectation about others' behavioral strategy for that stage. This expected payoff is used to calculate her logistic best response to arrive at the HAQRE-BU strategy profile.

Formally, the HAQRE-BU for finitely repeated games consists of a belief profile μ and a strategy profile b. Consider an arbitrary player-type's belief about the HQRE that is played in each stage $G^I(s)$ of Γ^I. Let this belief of λ_i^r about this strategy profile be denoted $\hat{b}_{-\lambda_i^r}$. Denote the belief of an arbitrary player-type λ_i^r about the likelihood that some opponent's type is λ_j^k, at an arbitrary stage s, following an arbitrary history a^{s-1} by $\mu_{\lambda_i^r}(\lambda_j^k|a^{s-1}, \hat{b}_{-\lambda_i^r}, \pi)$. Let the subjective expected payoff of the player-type λ_i^r in stage s from action $a_i \in A_i$ be given by $\bar{u}_{\lambda_i^r}(a_i \mid \hat{b}_{-\lambda_i^r}, \mu_{\lambda_i^r})$, and calculated by the standard method.

Definition 5 (HAQRE-BU for finitely repeated games). *For an arbitrary stage s of Γ^I, given an arbitrary history of play a^{s-1}, the belief of an arbitrary player-type λ_i^r about another player-type λ_j^k is given by the Bayes rule:*

$$\mu_{\lambda_i^r}(\lambda_j^k|a^{s-1}, \hat{b}_{-\lambda_i^r}, \pi) = \frac{Pr(\lambda_j^k \ and \ a^{s-1}|\hat{b}_{-\lambda_i^r}, \pi)}{Pr(a^{s-1}|\hat{b}_{-\lambda_i^r}, \pi)}. \tag{15}$$

And the logistic response of λ_i^r puts the following probabilities on each action $a_i \in A_i$:

$$b^{\lambda_i^r}(a_i|\hat{b}_{-\lambda_i^r}, \mu_{\lambda_i^r}) = \frac{exp[\lambda_i^r \bar{u}_{\lambda_i^r}(a_i \mid \hat{b}_{-\lambda_i^r}, \mu_{\lambda_i^r})]}{\sum_{a_i' \in A_i} exp[\lambda_i^r \bar{u}_{\lambda_i^r}(a_i' \mid \hat{b}_{-\lambda_i^r}, \mu_{\lambda_i^r})]}. \tag{16}$$

Due to the use of HQRE in each stage while forming beliefs and calculating expected payoffs, the existence of HAQRE-BU follows. But due to the possible multiplicity of HQRE, the uniqueness of HAQRE doesn't hold.

Proposition 5: *The HAQRE-BU for finitely repeated games exists.*

Due to finite repetitions, backward induction, and naive belief updating, in each stage a player-type maximizes payoffs only for that stage. A player-type does not need to consider any effects of his actions on subsequent behavior—an important consideration in the HAQRE (BU and non BU) for perfect-information games. Thus, in both versions of HAQRE for repeated games we don't need the recursive construction process. Note that, when there are multiple HAQRE, the expectation about others' strategies can be different for different player-types; i.e., $\hat{b}_{-\lambda_i^r}$ can be different from $\hat{b}_{-\lambda_j^k}$. Thus, the beliefs over a third player's types can be different across player-types at the same stage. This may provide too many degrees of freedom to researchers in games with multiple HAQRE. Thus, certain consistency conditions may be needed in such cases, e.g., that two players must agree on their belief about a third player's strategy.

Both notions of HAQRE, with and without BU, should be investigated in the important finitely-repeated-game context. Their data-fit can be compared with the data-fit of alternate behavioral models using appropriate empirical settings. By design, the HAQRE models should be most useful when there is a reason for heterogeneity in skill among the players. In what follows, we investigate the HAQRE and related models in the context of finite perfect information base games. The investigation of these models for finitely repeated games is left for future research.

6 Data application

The present work investigates whether accounting for heterogeneity in skill or payoff-responsiveness is important to model behavior in sequential-move games using QRE. For this data investigation, we focus on a base game of complete and perfect information. Specifically, we use the data from the incentivized Experiment 2 by Rampal (2020). This experiment was designed to study behavior in a sequential-move game in the presence of (induced) heterogeneity in skills across subjects and uncertainty about the opponent's skill-level.

Design: The experiment used an alternate-move two-player game called Avoid 13 and a variation of it called Computer 13. First, we focus on the Avoid 13 game. The first stage of the Avoid 13 game was used to decide the order of moves. Let us refer to the stages post the order-of-moves decisions as A13. In A13, the two players alternately choose to remove either 1, or 2, or 3 items from an imaginary box initially containing 13 items. The player who removes the last item(s) loses, and the opponent wins a prize. Therefore, Avoid 13 is a winner-take-all game. Backward induction provides a second-mover advantage in A13. In effect, a perfectly-playing second-mover can guarantee herself a win by removing, at each move, 4 minus the opponent's immediately-preceding choice. As illustrated in Figure 2, this procedure forces the opponent to remove the last item.

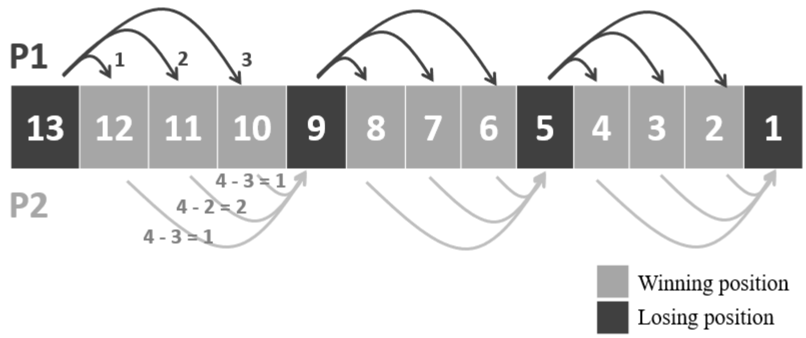

Figure 2: Second-mover's sure-win strategy in A13.

The prize for winning as first-mover is 500 experimental currency, and is larger than the prize for winning as second-mover: 200 experimental currency. To decide the order of moves in A13, the first stage of Avoid 13 is the "FS" stage: in this stage, both players simultaneously report whether they want to be the "First mover" (or F) or the "Second mover" (S); the decision of one of the players is selected at random and implemented. The decisions (henceforth "FS choices") remain private to each player. All subsequent moves in A13 are public information.

The Computer 13 game (C13 for short) is similar to Avoid 13, except for three key modifications. First, in C13 the human subject plays against a perfect-playing computer (i.e., the computer perfectly applies backward induction). Second, there is no extra incentive to win as the first-mover—there is a fixed prize for winning: 500 experimental currency. Last, the human player unilaterally decides whether to be the first or the second mover. Thus, a win in C13 is a strong signal that a player understands the relevant backward induction, and a loss in C13 is a clear signal that a player may have a weakness in the necessary backward induction. Each round in a session consisted of two "parts." First, each of the paired subjects played the C13 game against their respective computers, then they played the Avoid 13 game against each other.

The **heterogeneity** in skills across subjects was induced as follows. A session of the experiment contained between 8 and 18 subjects, who were split into two types: Experienced (Exp for short) and Inexperienced (Inexp). Each subject went through two sub-sessions. In the first sub-session (training sub-session), half the subjects were randomly selected to be made into Exp subjects; these subjects played 8 rounds (with random re-matching within the designated Exp subject pool) of the C13-Avoid 13 game. Meanwhile, Inexp subjects were kept unaware of Avoid 13 or C13, and instead they played an unrelated bargaining game. In the second sub-session (the combined sub-session), the Inexp subjects were mixed with the Exp subjects, and the entire subject pool for that session was randomly stranger re-matched into pairs at every round. The combined sub-session contained 6 rounds. Thus, an experimental session had 14 rounds in total.

The design also made **beliefs** important. In any round, the Exp or Inexp status of the opponent for that round was unknown. The overall proportion of Exp/Inexp subjects in each experimental session was publicly announced, about 50 percent each of Exp and Inexp. However, after the play of C13 by each subject and before the FS stage of Avoid 13, Exp subjects were privately given the opportunity to look at opponent's outcome (Win or Loss) and the latter's history of moves in C13. This opportunity was not given to Inexp subjects. Therefore, an Exp subject had the opportunity to update her beliefs about the opponent's type before the FS choice stage of Avoid 13. The Inexp subject could not update beliefs on the basis of the opponent's C13 history, but they could update beliefs on the basis of opponent's choices in A13. All aspects of the design were publicly announced to all subjects. See Rampal (2020) for more details.

6.1 Applicability of HAQRE and related models

We focus our analysis on the combined sub-session, where Exp and Inexp were mixed and randomly paired. No subject knew the Exp/Inexp status of the opponent in any round. But, they knew the proportions of such subjects in the population (the prior), and they had different degrees of information on the opponent's preceding choices.

Since half of the subjects were given more experience with the game in the training sub-session, we interpret differences in experience-level (Exp vs Inexp) as differences in skill. In effect, we assume that the two different levels-of-experience are associated with two different payoff responsiveness: λ_{Exp} being the common payoff-responsiveness parameter for experienced subjects, and analogously λ_{Inexp} for the inexperienced subjects. Thus, the heterogeneity in experience captures heterogeneity in skill or payoff responsiveness.

Since the number of subjects in each category (about 50 percent each) was publicly announced in each session, it seems appropriate to model the Experiment 2 combined sub-session setting as a common prior belief setting. Thus, we will assume that in each round, initially each subject believed that there was a 50 percent chance each that her opponent was Exp or Inexp.

Exp subjects could update beliefs on the basis of the opponent's play in C13. Both Exp and Inexp subjects observed all actions during the A13 stage of Avoid 13. Thus, both Exp and Inexp subjects had the opportunity to update beliefs about the skill-type of their opponent after each action of the latter. We account for all this possible belief updating (BU) in the BU versions of the models described earlier.

7 Results

There are two key aspects of the theory we want to invesitgate using data fit, as measured by maximized likelihood: is heterogeneity in skill important? Is belief-updating about the opponent's skill important? To answer the first question, we compare the AQRE with HAQRE; for the second question, we compare HAQRE with HAQRE-BU.

7.1 Likelihood comparison

We restrict our attention to the combined sub-session of Experiment 2 in Rampal (2020), where a difference in experience-level among subjects was induced, as described above. The data has 295 observations of rounds among different subject pairs. Each round had two parts: each pair of subjects first played C13 against respective computers, then played Avoid 13 against each other. An "outcome" of an experimental round is the combination of the win or loss outcome in C13 of the matched pair, the FS choice of each subject within the pair, and the sequence of moves taken by both subjects in Avoid 13 until the last item is removed.[9]

The optimal likelihood and the corresponding λ parameters of each model were obtained using maximum likelihood estimation (MLE). In the MLE exercise, we use the logit specification of the models. For AQRE, the MLE produces a single λ parameter, whereas for the HAQRE and HAQRE-BU models the optimal λ parameters are λ^{inexp} and λ^{exp}. In both HAQRE and HAQRE-BU, the common prior is $\pi = 0.5$. In the HAQRE-BU model, the belief of each subject is endogenously updated within a single play of the C13-Avoid 13 game, as per Definition 2 and the information available (as specified in the design).

Impact of heterogeneity

The AQRE model is nested in the HAQRE model (setting $\lambda^{inexp} = \lambda^{exp}$ in HAQRE yields the AQRE). Thus, to compare the data-fit of HAQRE vs AQRE using MLE, we use the Likelihood-Ratio test. We also compare the models in terms of the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC). The results and parameters are shown in Table 1.

The HAQRE model that accounts for heterogeneity yields a significantly better data-fit (p-value < 0.00001), with λ^{exp} estimated to be 0.015 and $\lambda^{inexp} = 0.009$. On the other hand, for AQRE, the homogeneous $\lambda = 0.011$. This provides evidence that incorporating heterogeneity in skill or payoff-responsiveness across players can be important for QRE-based models for sequential-move games. In this setting, the importance comes from the

[9] We do not account for possibly different sequences of moves in C13. We only make four categories of the observed outcome from C13—based on whether each player within the pair win or lose against the computer.

Results/Parameters	AQRE	HAQRE
LogLikelihood	-2453.44	-2403.32
AIC	4908.8	4810.6
BIC	4912.5	4818.0
λ_{inexp}	0.011649	0.009271
λ_{exp}		0.015214
p-value (H_0 : AQRE = HAQRE)	1.35×10^{-23}	

Table 1: MLE comparison between AQRE and HAQRE.

fact that players have heterogeneous experience levels, which can be a common occurrence in real-world interactions.

Within-the-play belief updating

Recall that in the experimental setting, Exp subjects had an additional opportunity to update beliefs: they observed the opponent's outcome (Win or Loss) against the computer (C13). Subsequently, both Exp and Inexp subjects could update their beliefs about the skill-type of the opponent by observing actions at each move of one's opponent during the A13 part of the game.

Let's illustrate the updating. Suppose that an Exp subject i holds a belief about the skill-types of her opponent j, which can be λ_j^{inexp} or λ_j^{exp}, and her prior belief is given by the commonly known prior distribution: 0.5 probability of each type. From the prize for winning in C13, she calculates the probability $Pr(\text{opponent loses in C13} \mid \text{opponent is Inexp})$ that her opponent loses in C13 conditional on being of type λ_j^{inexp}. She repeats this step for the type λ_j^{exp} of the opponent and uses the prior to calculate the unconditional probability $Pr(\text{opponent loses in C13})$. Thus, if the Exp subject observes that her opponent's outcome in C13 is Loss, she updates her beliefs about the opponent's skill-type using Bayes' rule as follows:

$$Pr(\text{opponent is Inexp} \mid \text{opponent lost in C13}) =$$
$$\frac{Pr(\text{opponent loses in C13} \mid \text{opponent is Inexp})\pi_j(\text{inexp})}{Pr(\text{opponent loses in C13})}.$$

In the HAQRE-BU model, we assume that all players update beliefs at every opportunity. The belief updating at other moves also follows the Bayes' rule. Take the experimental outcome LWSS323131: i.e., the First mover (in Avoid 13) lost in C13, the Second mover won in C13, both reported the intention to be "second mover", then players alternately removed 3, 2, 3, 1, 3, and 1 item(s). This outcome is particularly instructive of the mechanics of belief updating. Figure 3 illustrates that different magnitudes of updating are expected at different moments of the game. Let player 1 and player 2 stand for the selected first

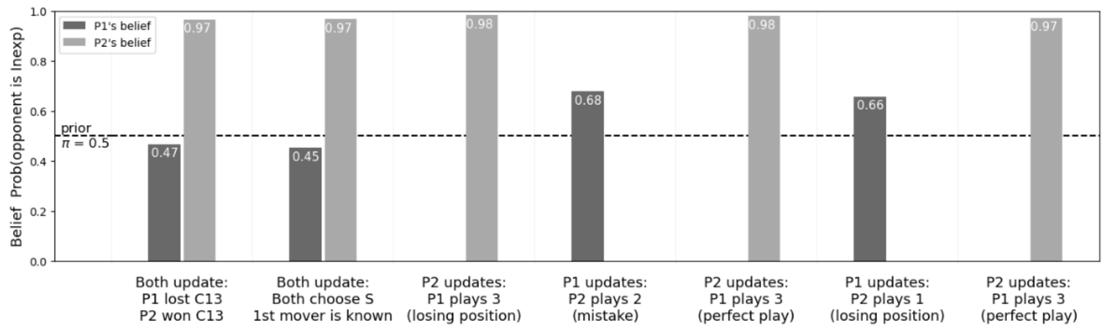

Figure 3: Belief profile corresponding to the outcome LWSS323131 among Exp players.

and second mover in Avoid 13, respectively. (i) After observing that the opponent lost against the perfect-playing computer, player 2's belief diverges radically from the prior—the updated belief puts almost probability 1 on player 1 being inexperienced. (ii) Decisions of the FS stage are not revealed, but one can update beliefs on the basis of one's own choice of F or S and the position one obtains.[10] (iii) When player 2 makes a mistake with 10 items left (a winning position), player 1's belief on player 2 being inexperienced increases substantially, as expected. (iv) Moves from losing positions or perfect-plays (from winning positions) have little impact on beliefs.

We use Vuong (1989) test for non-nested models to compare the optimal likelihoods of HAQRE and HAQRE-BU. The results and corresponding parameters are presented in Table 2, along with the AIC and BIC estimates.

Results/Parameters	HAQRE	HAQRE-BU
LogLikelihood	-2403.32	-2391.00
AIC	4810.6	4786.0
BIC	4818.0	4793.4
λ_{inexp}	0.009271	0.008925
λ_{exp}	0.015214	0.016196
p-value (H_0 : HAQRE = HAQRE-BU)	0.0268	

Table 2: MLE comparison between HAQRE and HAQRE-BU.

The HAQRE-BU model that allows for the dynamic belief updating by players about the opponent's payoff-responsiveness provides significant improvement in the data fit ($p-$value $= 0.027$). This result suggests that accounting for both heterogeneity in skills and belief-updating about the opponent's skill on the basis of preceding moves can be important to

[10]Thus, we see that player 1, who ends up in the first mover position despite choosing S updates slightly more than the second mover, player 2. Player 1 knows that the other must have chosen S; but player 2 cannot know whether his opponent chose F or whether the random assignment decided his second mover position.

explain behavior in sequential-move games using QRE. The importance of BU is demonstrated here in a setting where it is common knowledge that players with heterogeneous experience-levels are playing in the population.

Truncated and subjective beliefs

In this paper, we have also modeled subjective-prior formulations of HAQRE and HAQRE-BU. Recall that in the *Truncated* HAQRE model (THAQRE) only "downward-looking" beliefs are permitted. That is, player i with skill λ_i places 0 probability on types λ_j such that $\lambda_j > \lambda_i$ holds.

To model THAQRE in Rampal (2020)'s experimental setting, we assume that Inexp players presume that the opponent has the same payoff responsiveness as their own. Thus, an Inexp player i uses a degenerate probability distribution $\pi_j(\lambda_j^{inexp}) = 1$ when computing expected payoffs. Moreover, as described in the model of THAQRE, Inexp players assume that opponents also calculate expected payoffs using this degenerate distribution. On the other hand, an Exp player i still uses the prior distribution $\pi = 0.5$ on $\{\lambda_j^{exp}, \lambda_j^{inexp}\}$.

The THAQRE model yields improvement in data-fit when compared to the AQRE model ($p-$value $= 0.014$ in the Likelihood-Ratio test), which again corroborates to the importance of accounting for heterogeneity. On the other hand, THAQRE fits the data significantly worse than the HAQRE model ($p-$value $= 0.0003$ in Vuong (1989) test). This indicates that, in Rampal (2020)'s experiment, the publicly announced proportion of Exp/Inexp subjects is taken into account in prior-belief formation by Inexp players and prevails over the underlying reasons behind the hypothesis of truncated beliefs.

Subjective-priors can occur in ways other than "downward-looking" beliefs. In the Appendix, we apply a subjective-prior model where each experienced and inexperienced subject is sure that there is a 50-50 split of inexperienced and experienced subjects (as publicly announced), but experienced and inexperienced subjects have subjective beliefs about the values of the other type's payoff responsiveness. These are represented by $\hat{\lambda}^{inexp}$ (experienced types' common belief about an inexperienced type's payoff responsiveness) and $\hat{\lambda}^{exp}$ (inexperienced types' common belief about an experienced type's payoff responsiveness), respectively. Likelihood-Ratio tests for nested models show that SHAQRE and SHAQRE-BU fit the data significantly better than HAQRE and HAQRE-BU, respectively ($p-$values < 0.00001). But there is not much room for data-fit improvement left by incorporating belief updating once subjective priors are allowed: the difference between the optimal log-likelihoods of SHAQRE and SHAQRE-BU is not statistically significant. The details of this exercise are in the Appendix.

8 Conclusion

This paper extends the QRE model to sequential-move games where players can be of different skill types, and there is uncertainty and belief updating about each opponent's skill type. We define the Heterogeneous Agent Quantal Response Equilibrium (HAQRE) model for finite perfect information games and finitely repeated games, which extends the AQRE model of Mckelvey and Palfrey (1998) to account for skill heterogeneity, uncertainty, and naive belief updation. We also define variants of HAQRE with subjective and truncated prior beliefs, with and without naive belief updating. These definitions are provided in order to serve as modeling tools for applications where players with different degrees of skill interact in sequential-move games.

We show that each of these concepts exists, and in the finite perfect information game context, each of them generates a unique prediction. Furthermore, in the finite perfect information game context, the HAQRE and its variants do not require complicated fixed point calculations. Rather, a simple backward-induction-like recursive procedure yields the unique prediction given the prior on skill distributions assumed by the researcher, where skill is captured by payoff-responsiveness in QRE models.

We test the HAQRE and related models to experimental data of a winner-take-all sequential-move game between players drawn from two different experience-levels (high or low), with each player's own experience-level being private information. Results from our MLE exercises show that accounting for possible heterogeneity of skill can be important in such settings with possibly heterogeneous experience-level among players. The HAQRE model yields a significantly better data fit relative to AQRE. We also find that if, along with skill heterogeneity, we incorporate naive belief updating by players during the play of the sequential-move game, then that yields a further significant improvement in data fit. That is, we find that the HAQRE-BU model (HAQRE with belief updating) yields a significantly better data fit relative to the HAQRE.

Thus, we find that accounting for skill heterogeneity and belief updating about the opponent's skill, as we do in the HAQRE and related models, can be important in improving our understanding of multi-stage interactions among players with heterogeneous skill. The MLE parameters, yielding a greater payoff responsiveness estimate for experienced players, show that such situations of heterogeneity in skills can arise simply due to heterogeneity in experience—an arguable common occurrence in real world interactions.

Our data analysis deals with the context of finite perfect information games. Extending the data analysis of HAQRE and related models to the important context of finitely repeated games, especially when there is skill heterogeneity among players, is left for future research.

Appendix

A.1 HAQRE and HAQRE-BU for a general error distribution

General standard HAQRE

Definition 1 (HAQRE without belief updating) can be adapted to a general error distribution (which we call "general" for short, henceforth) using straightforward substitutions. Instead of Λ_i being the set of player-types for each player i of Γ, we have $\Phi_i = \{F_i^1, ..., F_i^{m(i)}\}$, where 1,.., $m(i)$ are natural numbers, and so $m(i)$ denotes the number of types of player i. Each player-type knows her type, but is uncertain about other players' types. This uncertainty is captured by a commonly-known prior distribution π, which specifies the prior over Φ_i for each player i in Γ. We assume that prior distributions across players are pairwise independent.

For each type $t \in \{1, .., m(i)\}$, F_i^t denotes the distribution from which each perceived payoff disturbance term (ϵ_{ika}^t) for player i of type t is drawn. Thus, ϵ_{ika}^t (used in (1)) is drawn independently and identically, according to the distribution F_i^t, for all $a \in A(h_i^k)$, and for all $h_i^k \in H^i$. We assume that the error distribution for each type of each player, F_i^t, is statistically independent from the error distribution for each type of each other player, $F_j^{t'}$, for all i, j and all t, t'. We will assume that for all player-types i-t, F_i^t with probability density function f_i^t is admissible in the sense of Mckelvey and Palfrey (1998). That is, we assume that each ϵ_i^t (the vector of all errors corresponding to all possible actions of player-type i-t) is an absolutely continuous random vector (with respect to the Lebesgue measure), and that the expected value of ϵ_{ika}^t exists for all i-t, h_i^k, and $a \in A(h_i^k)$.

Given the setup above, the definition of a general HAQRE b is exactly like the logistic HAQRE, except the following changes.

1. In step 1 of Definition 1, let $P(hS)$ be some player i, then the probability that each type F_i^t chooses an arbitrary action $a \in A(hS)$ is given by

$$b(a|hS, F_i^t) = Pr(a \textit{ yields } i \textit{ max expected payoff among } A(hS)|hS) =$$

$$= Pr[(\epsilon_{ihSa}^t)_{a \in A(hS)} | \overline{u}_i(a \mid hS) + \epsilon_{ihSa}^t \geq \overline{u}_i(a' \mid hS) + \epsilon_{ihSa'}^t \ \forall a' \in A(hS)]. \tag{17}$$

2. In step s, for $s \in \{2, .., S\}$ we need the following change. Type $F_{P(h_s)}^t$ of $P(h_s)$ can calculate the expected payoff of her action $a \in A(h_s)$ as:

$$\bar{u}_{P(h_s)}(a \mid h_s, b^s, \pi_{-P(h_s)}, F_{P(h_s)}^t).$$

Denote this as $\bar{u}(a|h_s)$ for short. Then, in the general HAQRE, the probability that each type $F^t_{P(h_s)}$ of $P(h_s)$ chooses an arbitrary action $a \in A(h_s)$ is given by

$$b(a|h_s, F^t_{P(h_s)}) =$$

$$= Pr[(\epsilon^t_{P(h_s)h_s a})_{a\in A(h_s)}|\bar{u}(a|h_s) + \epsilon^t_{P(h_s)h_s a} \geq \bar{u}(a'|h_s) + \epsilon^t_{P(h_s)h_s a'}\ \forall a' \in A(hS)]. \quad (18)$$

Given that the generalization of logistic HAQRE requires only these changes, and that (17) and (18) yield well defined and unique probabilities, we have Fact 1.

General HAQRE with belief updating

We now assume for HAQRE with belief updating (HAQRE-BU) the general error structure used for general (standard) HAQRE. We add one more assumption: for each player i and each type t, we only allow for F^t_i that yields a totally-mixed HAQRE.

The definition of a general HAQRE-BU uses the unique general HAQRE b^H, and (b, μ) is defined exactly like the logistic HAQRE-BU, except the following changes.

1. Beliefs are given as follows. For all $i \neq P(hs)$ and all $F^t_i \in \Phi_i$,

$$\mu_{F^t_{P(hs)}}(F^t_i|hs, b^H, \pi) = \frac{Pr(hs|\pi, b^H, F^t_i)}{\sum_{F^t_i\in\Phi_i} Pr(hs|\pi, b^H, F^t_i)}. \quad (19)$$

2. $F^t_{P(hs)}$'s expectation of payoffs from an arbitrary action $a \in A(hs)$ is:

$$\bar{u}_{P(hs)}(a \mid hs, b^H, \mu_{F^t_{P(hs)}}, b^{F^t_{P(hs)}}).$$

   Denote it as $\bar{u}(a|hs)$ for short. Then in HAQRE-BU, $F^t_{P(hs)}$ at her move hs chooses $a \in A(hs)$ with the probability

$$b(a|hs, F^t_{P(hs)}) =$$

$$= Pr[(\epsilon^t_{P(hs)hs\,a})_{a\in A(hs)}|\bar{u}(a|hs) + \epsilon^t_{P(hs)hs\,a} \geq \bar{u}(a'|hs) + \epsilon^t_{P(hs)hs\,a'}\ \forall a' \in A(hs)]. \quad (20)$$

3. The construction of $b^{F^t_{P(hs)}}$ is the same as the construction of $b^{\lambda_{P(hs)}}$ (other than μ and b^H different as specified above), except that to calculate the probabilities with which a certain action is to be played at any given move, $F^t_{P(hs)}$ is used rather than the logit formulation used there.

Given that the generalization of logistic HAQRE-BU requires only these changes, and that (19) and (20) yield well defined and unique probabilities, we have Fact 2.

A.2 Subjective HAQRE

Recall that we used Γ^I to denote the incomplete-information game with observable actions, where for each player $i \in N$ of Γ, there is a set of player-types Λ_i, and there is a common knowledge prior probability distribution over the types of each player i given by π_i for $i \in N$. To specify the Subjective HAQRE model (SHAQRE for short), we will need an additional object in Γ^I: each player-type t_i (type t of player i) will have a subjective prior belief given by ρ_{t_i}, where ρ_{t_i} specifies a probability distribution over $\times_{i\in N}\Lambda_i$. When we refer to Γ^I in this section, it is understood that Γ^I includes the subjective prior beliefs for each player type.

Subjective HAQRE without belief updating

The logistic version of SHAQRE (this is without belief updating) is a behavioral strategy profile for Γ^I which specifies for each type of each player, the (possibly mixed) action they will take at each move. There is only one change required relative to the Logistic HAQRE to define the Logistic SHAQRE. In steps 2 through S of the HAQRE definition (Definition 1), we replace "*the probability distributions over other players' types,* $\pi_{-P(h_s)}$," used by $\lambda_{P(h_s)}$ to calculate the expected payoff from each action in $A(h_s)$, by the subjective prior of $\lambda_{P(h_s)}$ which we can denote by $\rho_{-\lambda_{P(h_s)}}$. Note that each player-type uses the same subjective prior at each of her moves, i.e., $\rho_{-\lambda_{P(h_s)}}=\rho_{-\lambda_{P(h_s')}}$ whenever $\lambda_{P(h_s)} = \lambda_{P(h_s')}$, even if $h_s \neq h_s'$. Thus, analogous to Proposition 1, existence and uniqueness follows for the Logistic SHAQRE.

For the case of general error structures, we replace Λ_i as the set of player-types for each player i of Γ^I, and instead we have $\Phi_i = \{F_i^1, ..., F_i^{m(i)}\}$ as the set of player-types. We add one more assumption: we only allow for F_i^t for each player i and each type t that yields a totally-mixed SHAQRE. Again, instead of the commonly-known prior distribution π, each type of each player has a different prior belief given by $\rho_{F_i^t}$. So, the definition of a general SHAQRE b is exactly like the general HAQRE, except that in (18) we use $\rho_{-F^t_{P(h_s)}}$ instead of $\pi_{-P(h_s)}$ to calculate expected payoffs from actions available at $A(h_s)$. Again, analogous to Fact 1, for every Γ^I, the general SHAQRE exists, and it is unique.

Subjective HAQRE with belief updating

The belief updating in the Logistic SHAQRE-BU is exactly as it is in Logistic HAQRE-BU, except the prior belief of each player-type is subjective and not the common-knowledge objective prior. Again, we will assume naive belief updating: i.e., we will assume that while player-types update beliefs about opponents' types on the basis of opponents' actions they observe, they don't account for opponents' types performing the analogous belief updating.

Instead, each player-type t_i assumes that each type of each opponent, say t'_j holds prior beliefs which are the same as the subjective prior beliefs of t_i, denoted as ρ_{t_i}, at each of the moves of t'_j. Again, the rationale for this choice is incorporating simplicity in the reasoning process of each player-type.

First, we describe how the Logistic HAQRE-BU definition (Definition 2) needs to be modified to describe the Logistic SHAQRE-BU. The Logistic SHAQRE with belief updating (SHAQRE-BU) is a behavioral strategy profile b and a profile of beliefs $\mu = (\mu_{\lambda_i})_{\lambda_i \in \Lambda_i \text{ and } i \in N}$ for Γ^I. We first define μ. Consider the unique SHAQRE strategy profile of Γ^I, denoted b^H. For any information set hs with some player-type $\lambda_{P(hs)}$ moving there, the SHAQRE-BU belief of $\lambda_{P(hs)}$ is calculated through Bayesian updating using the subjective prior distribution of $\rho_{\lambda_{P(hs)}}$ and the SHAQRE strategy profile b^H. Thus, we have for all $i \neq P(hs)$ and all $\lambda_i \in \Lambda_i$,

$$\mu_{\lambda_{P(hs)}}(\lambda_i | hs, b^H, \rho_{\lambda_{P(hs)}}) = \frac{Pr(hs|\rho_{\lambda_{P(hs)}}, b^H, \lambda_i)}{\sum_{\lambda'_i \in \Lambda_i} Pr(hs|\rho_{\lambda_{P(hs)}}, b^H, \lambda'_i)}. \tag{21}$$

Given that μ for SHAQRE-BU is defined as in (21), the construction of the strategy profile b of the Logistic SHAQRE-BU using μ is exactly as in Definition 2. Furthermore, analogous to Proposition 2, the Logistic SHAQRE-BU exists, and it is unique.

SHAQRE-BU for a general error structure

To model SHAQRE with belief updating for general error structures (as in general SHAQRE), we again add one more assumption: we only allow for F_i^t for each player i and each type t that yields a totally-mixed SHAQRE. The subjective prior of F_i^t is given by $\pi_{F_i^t}$.

The definition of a general SHAQRE-BU uses the unique general SHAQRE b^H of Γ^I, and (b, μ) is defined exactly like the Logistic SHAQRE-BU, except the following change. Beliefs are calculated using subjective priors rather than an objective common knowledge prior. So, for all $i \neq P(hs)$ and all $F_i^t \in \Phi_i$, we have

$$\mu_{F^t_{P(hs)}}(F_i^t | hs, b^H, \rho_{F^t_{P(hs)}}) = \frac{Pr(hs|\rho_{F^t_{P(hs)}}, b^H, F_i^t)}{\sum_{F_i^t \in \Phi_i} Pr(hs|\rho_{F^t_{P(hs)}}, b^H, F_i^t)}. \tag{22}$$

Finally, the strategy profile b of the general SHAQRE-BU is constructed from μ as in (20). Given that the generalization of the Logistic SHAQRE-BU requires only these changes, and that (22) and the associated b yield well defined and unique probabilities, analogous to Fact 2, the general SHAQRE-BU exists, and it is unique.

A.3 MLE results for logit-SHAQRE and logit-SHAQRE-BU

Here we detail the results obtained for the logistic specification of SHAQRE and SHAQRE-BU with respect to the experimental data in Rampal (2020). The MLE exercise for subjective models in this setting gives a set of optimal $\lambda-$parameters composed by four elements: the two real payoff-responsiveness of Exp and Inexp subjects, respectively, λ^{exp} and λ^{inexp}, and each type's common belief about the other type's payoff responsiveness, $\hat{\lambda}^{exp}$ and $\hat{\lambda}^{inexp}$.

The maximized likelihoods, the AIC and BIC estimates and the optimal parameters for SHAQRE and SHAQRE-BU are presented in Table 3 next to their respective (non-subjective) counterparts. Table 3 also shows the results of the Likelihood-Ratio tests comparing each subjective model to its counterpart.

Results/Parameters	HAQRE	SHAQRE	HAQRE-BU	SHAQRE-BU
LogLikelihood	-2403.32	-2261.72	-2391.00	-2270.51
AIC	4810.6	4531.4	4786.0	4549.0
BIC	4818.0	4546.1	4793.4	4563.7
$\lambda^{inexp} \mid \hat{\lambda}^{inexp}$	0.009271	0.0070 | 0.0270	0.008925	0.0075 | 0.0278
$\lambda^{exp} \mid \hat{\lambda}^{exp}$	0.015214	0.0095 | 0.0393	0.016196	0.0090 | 0.0165
p-value (H_0 : HAQRE(-BU) = SHAQRE(-BU))	$3.20 \cdot 10^{-62}$		$4.68 \cdot 10^{-53}$	

Table 3: MLE Comparison: HAQRE vs SHAQRE, BU and non-BU.

The SHAQRE and SHAQRE-BU models fit the data significantly better than HAQRE and HAQRE-BU, respectively ($p-$values < 0.00001). On the other hand, the difference between the optimal likelihoods of SHAQRE and SHAQRE-BU is not statistically significant ($p-$value $= 0.20$). This indicates that most of the significant data-fit improvement is incorporated into the priors once subjectivity is allowed, and there is not much room left for improvement coming from belief updating.

References

Anderson, L. R. and C. A. Holt (1997). Information cascades in the laboratory. *American Economic Review (87)*, 847–862.

Breitmoser, Y., J. H. W. Tan, and D. J. Zizzo (2014). On the beliefs off the path: Equilibrium refinement due to quantal response and level-k. *Games and Economic Behavior (86)*, 102–125.

Camerer, C. and D. Lovallo (1999). Overconfidence and excess entry: An experimental approach. *American Economic Review (89)*, 306–318.

Camerer, C., S. Nunnari, and T. R. Palfrey (2016). Quantal response and nonequilibrium beliefs explain overbidding in maximum-value auctions. *Games and Economic Behavior (98)*, 243–263.

Camerer, C. F., T.-H. Ho, and J.-K. Chong (2004). A cognitive hierarchy model of games. *Quarterly Journal of Economics (119)*, 861–898.

Carrillo, J. D. and T. R. Palfrey (2009). The compromise game: two-sided adverse selection in the laboratory. *American Economic Journal: Microeconomics (1)*, 151–181.

Chen, Y., X. Su, and X. Zhao (2012). Modeling bounded rationality in capacity allocation games with the quantal response equilibrium. *Management Science (58)*, 1952–1962.

Cui, T. H. and Y. Zhang (2018). Cognitive hierarchy in capacity allocation games. *Management Science (64)*, 1250–1270.

Friedman, E. (2020). Endogenous quantal response equilibrium. *Games and Economic Behavior (124)*, 620–643.

Friedman, E. (2022). Stochastic equilibria: Noise in actions or beliefs? *American Economic Journal: Microeconomics (14)*, 94–142.

Goeree, J. K. and C. A. Holt (2004). A model of noisy introspection. *Games and Economic Behavior (46)*, 365–382.

Goeree, J. K., C. A. Holt, and T. R. Palfrey (2016). *A Stochastic Theory of Games*. Princeton University Press.

Goeree, J. K., P. Louis, and J. Zhang (2018). Noisy introspection in the 11-20 game. *Economic Journal (128)*, 1509–1530.

Ho, T.-H. and X. Su (2013). A dynamic level- <i>k</i> model in sequential games. *Management Science (59)*, 452–469.

Ivanov, A., D. Levin, and J. Peck (2009). Hindsight, foresight, and insight: An experimental study of a small-market investment game with common and private values. *American Economic Review (99)*, 1484–1507.

Kahneman, D. and A. Tversky (1973). On the psychology of prediction. *Psychological Review (80)*, 237–251.

Kreps, D. M., P. Milgrom, J. Roberts, and R. Wilson (1982). Rational cooperation in the finitely repeated prisoners' dilemma. *Journal of Economic Theory (27)*, 245–252.

Kubler, D. and G. Weizsacker (2004). Limited depth of reasoning and failure of cascade formation in the laboratory. *Review of Economic Studies (71)*, 425–441.

McKelvey, R. D. and T. R. Palfrey (1995). Quantal response equilibria for normal form games. *Games and Economic Behavior (10)*, 6–38.

Mckelvey, R. D. and T. R. Palfrey (1998). Quantal response equilibria for extensive form games. *Experimental Economics (1)*, 9–41.

Mengel, F. (2014). Learning by (limited) forward looking players. *Journal of Economic Behavior {&} Organization (108)*, 59–77.

Moinas, S. and S. Pouget (2013). The bubble game: An experimental study of speculation. *Econometrica (81)*, 1507–1539.

Nyarko, Y. and A. Schotter (2002). An experimental study of belief learning using elicited beliefs. *Econometrica (70)*, 971–1005.

Osborne, M. J. and A. Rubinstein (1994). *A course in game theory*. MIT Press.

Rampal, J. (2020). Opponent's foresight and optimal choices.

Rogers, B. W., T. R. Palfrey, and C. F. Camerer (2009). Heterogeneous quantal response equilibrium and cognitive hierarchies. *Journal of Economic Theory (144)*, 1440–1467.

Tingley, D. and S. Wang (2010). Belief updating in sequential games of two-sided incomplete information: An experimental study of a crisis bargaining model. *Quarterly Journal of Political Science (5)*, 243–255.

Vuong, Q. H. (1989). Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica (57)*, 307.