# Cognitive Hierarchies in Multi-Stage Games of Incomplete Information: Theory and Experiment*

Po-Hsuan Lin†

October 22, 2025

**Abstract**

This paper theoretically and experimentally studies the dynamic cognitive hierarchy (DCH) solution, a novel hierarchical reasoning model for general extensive games, in multi-stage games of incomplete information. Using the dirty-faces game, I illustrate how the DCH solution can differ dramatically between strategically equivalent games. I test this prediction in a laboratory experiment using two strategically equivalent versions of the dirty-faces game. The parameters are calibrated to maximize the informativeness of the experiment. The experimental results reveal significant differences in behavior between the two versions, and more importantly, the observed differences align with DCH. This result suggests playing a dynamic game in a different, but strategically equivalent, version can lead to distortions in behavior in a predictable way.

**JEL Classification Numbers:** C72, C92, D82, D83

**Keywords:** Level-$k$, Dirty-Faces Game, Learning, Strategic Equivalence, Strategy Method

---

*This paper was previously circulated under the title "Cognitive Hierarchies in Multi-Stage Games of Incomplete Information." I am indebted to my advisor, Thomas Palfrey, for his constant support and encouragement. I thank Marina Agranov, Pierpaolo Battigalli, Colin Camerer, Anujit Chakraborty, John Duffy, Michael Gibilisco, Charles Holt, Paul J. Healy, Greg Leo, Rosemarie Nagel, Kirby Nielsen, Jean-Laurent Rosenthal, Burkhard C. Schipper, Charles Sprenger, Omer Tamuz, Joseph Tao-yi Wang, Yi Xin and audiences at California Institute of Technology, Stanford University, UC Davis, University of Houston, University of Virginia, Virginia Commonwealth University, the 2022 ESA North American Conference, the 2023 ESA North American Conference, the 2023 Los Angeles Experiments (LAX) Workshop and BEEMA9 for their valuable comments. I also thank John Duffy and Michael McBride for their generosity and support of my use of the Experimental Social Science Laboratory at UC Irvine. Lastly, I am grateful to Roberto Weber and Ralph Bayer for sharing their data. The experiment is approved by Caltech IRB #22-1256. This work is supported by NSF Dissertation Improvement Grant #2243268. All errors are my own.

†Department of Economics, University of Virginia, Charlottesville, VA 22904. Email: plin@virginia.edu

# 1 Introduction

In many economic settings, people interact with each other over time under asymmetric information, updating their beliefs about other people's private information based on observed actions—such as signaling, reputation building, bargaining, and communication. In game theory, the standard approach to analyzing these situations is to model them as multi-stage games of incomplete information and solve for the sequential equilibrium, assuming that players sequentially best respond to their beliefs at every information set.

However, the best response requirement alone cannot pin down precise predictions because players may act rationally given their beliefs but may hold mistaken beliefs about what others will do. Consequently, equilibrium theory imposes a strong assumption about strategic sophistication, requiring players to correctly anticipate the actions of others in equilibrium—a requirement known as *mutual consistency*. From an empirical standpoint, this requirement may be implausibly strong, particularly in complex multi-stage games of incomplete information, as indicated by behavior observed in many laboratory experiments (see, for example, Camerer, 2003).

To illustrate how demanding the mutual consistency requirement is in the context of multi-stage games of incomplete information, consider a simple Dirty-Faces Game.[1]

> This game involves two players, Ann and Bob. Their faces are either "clean" or "dirty," and each player can see the other's face but not their own. If at least one player has a dirty face, a public announcement is made to ensure that this information becomes common knowledge.
>
> After observing the other's face and the announcement, players take actions over a series of periods. In each period, players simultaneously choose between "wait" and "claim" (to have a dirty face). The actions are revealed to both players at the end of each period. If both players decide to wait, the game proceeds to the next period. Otherwise, the game ends after the period in which at least one player chooses to claim.
>
> Players receive rewards for correctly claiming to have a dirty face but are penalized for making false claims. The payoffs are structured such that players should

[1]The original dirty-faces game is introduced by Littlewood (1953) as a three-player logical puzzle: "Three ladies, A,B,C, in a railway carriage all have dirty faces and are all laughing. It suddenly flashes on A: why doesn't B realize C is laughing at her? Heavens! I must be laughable." (pp. 3–4 in Littlewood 1953) Additionally, the dirty-faces game has also been referred to as the "cheating wives puzzle" (Gamow and Stern, 1958), the "cheating husbands puzzle" (Moses et al., 1986), the "muddy children puzzle" (Barwise, 1981; Halpern and Moses, 1990), and the "red hat puzzle" (Hardin and Taylor, 2008).

"wait" if they are uncertain about their face but "claim" as soon as they know their face is dirty.[2]

This (two-person) dirty-faces game is arguably one of the simplest multi-stage games of incomplete information, as the unique equilibrium can be solved using a straightforward two-step reasoning: When a player, say Ann, sees a clean face (along with the announcement), rational Ann should recognize that her own face is dirty and claim in period 1. However, if rational Ann sees a dirty face, she will wait in period 1 since she has no knowledge of her own face type. In equilibrium, Ann believes that *rational Bob would have claimed in the first period if her own face were clean.* Thus, if the game does indeed reach the second period, Ann will realize that her face is dirty and claim.

In other words, standard equilibrium theory makes the bold prediction that there will be no asymmetric information after period 1, as long as players are rational and believe that the other player is also rational. Nevertheless, despite the simplicity of the game's structure and the clarity of the equilibrium reasoning, the first dirty-faces game experiment conducted by Weber (2001) revealed a significant inconsistency between observed behavior and the equilibrium prediction. Specifically, Weber (2001) found that half of the subjects failed to perform this two-step reasoning, highlighting how behaviorally demanding mutual consistency can be—particularly in multi-stage games with incomplete information.[3]

In contrast to the standard equilibrium theory, this paper theoretically and experimentally studies the "Dynamic Cognitive Hierarchy (DCH) Solution," a non-equilibrium behavioral game theory for general extensive games developed by Lin and Palfrey (2024), within the context of multi-stage games of incomplete information. The DCH solution is in the same spirit of the level-$k$ model (Nagel, 1995) and the Cognitive Hierarchy (CH) theory (Camerer et al., 2004), in which players are endowed with heterogeneous levels of sophistication and best respond to the mistaken belief that other players are less sophisticated than themselves.

Specifically, the DCH solution assumes a hierarchical structure of strategic sophistication among players. Level-0 players are non-strategic and uniformly randomize at each information set. Level-$k$ players, on the other hand, mistakenly believe that all other players are distributed across levels ranging from 0 to $k-1$ and (sequentially) best respond to this incorrect belief.[4] The DCH solution differs from the standard level-$k$/CH framework in that it is

[2]See Section 3 for a detailed description of the dirty-faces game.

[3]As reviewed by Camerer (2003), "Weber's results are important because Caltech students are selected for their skill at logic puzzles such as the dirty-faces game (and they know others students are very skilled). The extend of their iterated reasoning—half of them do not do two steps—is a plausible upper bound on iterative reasoning by most people in abstract games such as these." (pp. 238–239 in Camerer 2003)

[4]In this paper, level-0 players may be interchangeably referred to as "non-strategic players," and level-$k \geq 1$ players as "strategic players" since they best respond to their beliefs.

defined with respect to the extensive form of a game rather than the normal form, allowing players to update their beliefs about other players' levels upon reaching different information sets. In multi-stage games of incomplete information, DCH posits that players will update their beliefs about both other players' levels *and* payoff-relevant private types. This makes the learning dynamics of DCH fundamentally different from the standard equilibrium theory, allowing DCH to generate more behaviorally plausible predictions. In this paper, I use the dirty-faces game to illustrate the sharp difference between DCH and equilibrium theory.

DCH predicts that in the first period of the dirty-faces game, all strategic players behave like rational players—claiming immediately upon seeing a clean face and the announcement, while waiting when they see a dirty face. In contrast, level-0 players choose randomly, regardless of their observations. Unlike strategic players, *the actions of level-0 players convey no information about the true face types.* Therefore, when observing a dirty face and the game proceeds beyond period 1, strategic players will believe that there are two possible situations:

1. If the other player is level-0, then their action is randomly determined and provides no information about my own face.

2. If the other player is level-$k \geq 1$, then my own face is certainly dirty because the other player would have claimed in period 1 if my own face were clean.

Consequently, after period 1, a strategic player faces a dynamic tradeoff. If they wait and the game proceeds to the next period, they will become more certain about having a dirty face because the other player is less likely to be a level-0 player. However, the risk of waiting is that the game might be randomly terminated (by a level-0 opponent) and the payoff is further discounted due to impatience. In other words, DCH posits that strategic players essentially face a dynamic programming problem, with the DCH solution characterized by level-dependent stopping periods, which depend on the prior distribution of levels and the payoffs parameters.

Besides the dramatic difference in learning dynamics, this paper also highlights an important property of DCH—*the violation of invariance under strategic equivalence.* Two extensive games are *strategically equivalent* if they share the same reduced normal form. Since the DCH solution is defined with respect to the extensive form of a game, it can generally differ between two strategically equivalent extensive games if their extensive forms differ.

From the perspective of standard game theory, implementing a game in its extensive form or reduced normal form is considered equivalent, as the set of Nash equilibrium outcomes is *reduced normal form invariant.* That is, if two extensive games share the same reduced normal form, they will have the same set of Nash equilibria.[5] However, from the perspective of DCH, even if two extensive games share the same reduced normal form, the size of action sets may differ when outcome-equivalent strategies are consolidated into reduced strategies. This difference leads non-strategic level-0 players to behave differently, which in turn influences all higher-level players. Since DCH is solved recursively from the bottom of the hierarchy, these differences ultimately result in distinct DCH solutions across strategically equivalent games.

The concept of strategic equivalence is crucial and widely applied in both mechanism design and experimental methodology. In many institutional settings, the reduced normal form mechanism is often implemented instead of the extensive form mechanism due to simplicity or practicality. However, growing evidence from various contexts suggests that strategically equivalent games, when represented differently, can lead to different behaviors—for example, sealed-bid auctions versus clock auctions (Li, 2017), runoff elections versus contingent elections (Richie et al., 2023), and dynamic versus static matching mechanisms (Gong and Liang, 2024). Additionally, when conducting experiments on extensive games, experimenters often adopt the "strategy method" proposed by Selten (1967), which implements the game in its reduced normal form, rather than the "direct-response method," which implements the game in its extensive form. The primary advantage of the strategy method is its ability to collect a larger amount of data, particularly at information sets that are rarely reached in actual play. Yet whether the use of strategy method introduces behavioral distortions remains an ongoing debate in experimental methodology (see, for example, Brandts and Charness 2011; García-Pola et al. 2020). In this regard, DCH emerges as a theory that can potentially predict both the occurrence and direction of violations of invariance under strategic equivalence.

To empirically test the violation of invariance under strategic equivalence as predicted by DCH, I design and conduct a dirty-faces game experiment with two treatments that implement strategically equivalent versions of the game. In the *Sequential* treatment, the game is played period by period, allowing players to observe the history and make decisions at realized information sets. In contrast, the *Simultaneous* treatment presents the game in its reduced normal form, where players select their contingent strategies. While equilibrium theory predicts no difference between the two treatments, DCH predicts a treatment effect, with the predicted magnitude depending on both the game parameters and the true distribution of levels, which is unknown prior to conducting the experiment.

In this case, how should we design the payoff parameters? To address this, I develop an

[5] In fact, it has been argued that the property of reduced normal form invariance is a desired requirement for any set valued solution concept for noncooperative games (Kohlberg and Mertens, 1986).

*optimal design approach*, in which I first estimate the distribution of levels using data from a dirty-faces game experiment reported by Bayer and Chan (2007).[6] Then, I select game parameters that maximize the informativeness of the experiment by considering a mix of parameters expected to produce various magnitudes of the treatment effect.

By employing this fine-tuned experimental design, significant treatment effects are detected in the data. More importantly, *how* the invariance is violated aligns with the predictions of DCH. Additionally, to examine whether the observed differences in the data can be attributed to the violation of other equilibrium requirements, I compare DCH with two alternative models: Agent Quantal Response Equilibrium by McKelvey and Palfrey (1998) and Cursed Sequential Equilibrium by Fong et al. (2025). These alternative models relax the best response requirement and Bayesian inference, respectively. While there is evidence of the failure of best responses and Bayesian inference, the observed violation of invariance in the data is primarily driven by the violation of mutual consistency, as DCH significantly outperforms the alternative models in both treatments. These experimental results highlight the behavioral fragility of invariance under strategic equivalence in multi-stage games of incomplete information and demonstrate that DCH provides a more reasonable evaluation of behavioral distortions across different representations.

**Related Literature.** This paper builds on the extensive literature on the level-$k$ model and the cognitive hierarchy theory. Since its introduction by Nagel (1995) and Stahl and Wilson (1994, 1995), the level-$k$ model has been widely used to analyze a variety of laboratory experimental data, including the beauty contest (guessing) game (Nagel, 1995; Costa-Gomes and Crawford, 2006), simultaneous-move matrix games (Costa-Gomes et al., 2001; Crawford and Iriberri, 2007a), the market entry game (Camerer et al., 2004), auctions (Crawford and Iriberri, 2007b), the trading game (Rogers et al., 2009) and the sender-receiver cheap talk game (Cai and Wang, 2006).

In addition to laboratory experimental data, the level-$k$/CH model has also been applied to explain non-equilibrium behavior in the field, such as US timber auctions (Gillen, 2009; An, 2017), telephone market entry (Goldfarb and Xiao, 2011), the Swedish LUPI lottery (Östling et al., 2011), movie reviews (Brown et al., 2012), the Texas spot electricity market (Hortaçsu et al., 2019) and the forward guidance puzzle in macroeconomics (Bersson et al., 2024). Furthermore, the rationale behind this hierarchical reasoning model is supported by neuropsychological evidence from fMRI studies (Coricelli and Nagel, 2009) and eye-tracking experiments (Wang et al., 2010).

[6] Bayer and Chan (2007) later replicated Weber (2001) in a more generalized class of dirty-faces games with a larger dataset, making Bayer and Chan (2007) a more suitable dataset for calibration purposes.

Despite its success in explaining non-equilibrium behavior in both laboratory and the field, a major limitation of this approach is that the standard level-$k$/CH model is defined only for simultaneous-move games. To address this limitation, Lin and Palfrey (2024) propose the Dynamic Cognitive Hierarchy (DCH) solution, which extends the CH framework to general extensive games[7] and highlights the theoretical implications of the violation of invariance under strategic equivalence in games of perfect information. Building on this work, this paper further explores, both theoretically and experimentally, the violation of invariance under strategic equivalence in multi-stage games of incomplete information.

At a more conceptual level, DCH is related to other behavioral solution concepts for dynamic games that relax different equilibrium requirements. DCH is a non-equilibrium model in which different levels of players best respond to different conjectures about other players' strategies. In contrast, Agent Quantal Response Equilibrium (AQRE), introduced by McKelvey and Palfrey (1998), is an equilibrium model where players make stochastic choices. Both DCH and AQRE assume that players use Bayes' rule to update their beliefs. On the other hand, Cursed Sequential Equilibrium (CSE), proposed by Fong et al. (2025), is an equilibrium model in which players are able to make best responses but fail to make correct Bayesian inference, thereby capturing scenarios where players do not fully understand how other players' actions depend on their private information.[8] A common theoretical property shared by these behavioral solution concepts is the violation of invariance under strategic equivalence, albeit in different ways. To gain insights into how these models contribute to the observed violation of invariance under strategic equivalence, I structurally estimate each model and compare their respective goodness of fit.

Finally, this paper contributes to the literature on dirty-faces game experiments. Weber (2001) and Bayer and Chan (2007) conducted the first experiments on the dirty-faces game, finding that many subjects failed to engage in iterative reasoning. Recent studies have further demonstrated the persistence of this failure, even when subjects play against fully rational robot opponents (Grehl and Tutić, 2015; Chen et al., 2024). Other research has also shown that the failure of iterative reasoning is correlated with cognitive abilities (Devetag and Warglien, 2003; Bayer and Renou, 2016a,b), while deviations from equilibrium decrease when participants are selected through a market mechanism (Choo and Zhou, 2022). Taken together, these experimental findings suggest substantial heterogeneity in strategic sophistication within the population.

[7] It is worth noting that Schipper and Zhou (2024) contemporaneously propose an alternative extension of the standard level-$k$ model to extensive games by allowing level-$k$ players to update their beliefs in a non-Bayesian way, whereas DCH posits that strategic levels of players update their beliefs using Bayes' rule.

[8] Sequential Cursed Equilibrium (SCE), proposed by Cohen and Li (2022), is an alternative model that captures the bias in which individuals fail to recognize how others' actions depend on their information set partitions. For a detailed comparison of CSE and SCE, see Fong et al. (2023).

**Outline.** The rest of the paper is organized as follows. Section 2 introduces the DCH solution and establishes its general properties in multi-stage games of incomplete information. In Section 3, I analyze the DCH solution in two different but strategically equivalent versions of the dirty-faces game. Section 4 outlines the experimental design, and Section 5 presents the experimental results. Finally, Section 6 concludes the paper.

# 2 Model

Section 2.1 introduces the framework for multi-stage games of incomplete information, following Fudenberg and Tirole (1991). The DCH solution for this class of games is defined in Section 2.2. Finally, Section 2.3 establishes several general properties of DCH in multi-stage games of incomplete information.

## 2.1 Multi-Stage Games of Incomplete Information

Let $N = \{1, \ldots, n\}$ be a finite set of players. Each player $i \in N$ has a payoff-relevant private *type* $\theta_i$ drawn from a finite set $\Theta_i$. Let $\theta \in \Theta \equiv \times_{i=1}^{n} \Theta_i$ represent the type profile and $\theta_{-i}$ be the type profile without player $i$. All players have the common (full support) prior distribution $F : \Theta \to (0, 1)$. At the beginning of the game, each player is told their own type, but is not informed anything about the types of others. Therefore, each player $i$'s initial belief about the types of others when their type is $\theta_i$ is:

$$F(\theta_{-i}|\theta_i) = \frac{F(\theta_{-i}, \theta_i)}{\sum_{\theta'_{-i} \in \Theta_{-i}} F(\theta'_{-i}, \theta_i)}.$$

If the types are independent across players, each player $i$'s initial belief about the types of others is $F_{-i}(\theta_{-i}) = \Pi_{j \neq i} F_j(\theta_j)$ where $F_j(\theta_j)$ is the marginal distribution of player $j$'s type.

The game is played in "periods" $t = 1, 2, \ldots, T$ where $T < \infty$. In each period, players simultaneously choose their actions, which will be revealed at the end of the period. The feasible set of actions can vary with histories, so games with alternating moves are also included. Let $\mathcal{H}^{t-1}$ be the set of all available histories at period $t$, where $\mathcal{H}^0 = \{h_\emptyset\}$ and $\mathcal{H}^T$ is the set of terminal histories. Let $\mathcal{H} = \bigcup_{t=0}^{T} \mathcal{H}^t$ be the set of all available histories of the game, and let $\mathcal{H} \backslash \mathcal{H}^T$ be the set of non-terminal histories.

For every player $i$, since the available information at any period $t$ is the history $h^{t-1}$ and his own type $\theta_i$, the information set corresponds to player $i$'s move at period $t$ is identified with an element of $\Theta_i \times \mathcal{H}^{t-1} \equiv \mathcal{I}_i$. For the sake of simplicity, I assume that, at each history, the feasible set of actions for every player is independent of their type and use $A_i(h^{t-1})$ to denote the feasible set of actions for player $i$ at history $h^{t-1}$. Let $A_i = \bigcup_{h \in \mathcal{H} \setminus \mathcal{H}^T} A_i(h)$ denote player $i$'s feasible actions in all histories of the game and $A = A_1 \times \cdots \times A_n$. In addition, I assume $A_i$ is finite for all $i \in N$ and $|A_i(h)| \geq 1$ for all $i \in N$ and any $h \in \mathcal{H} \setminus \mathcal{H}^T$.

A behavioral strategy for player $i$ is a function $\sigma_i : \mathcal{I}_i \to \Delta(A_i)$ satisfying $\sigma_i(\theta_i, h^{t-1}) \in \Delta A_i(h^{t-1})$. Furthermore, I use $\sigma_i(a_i^t | \theta_i, h^{t-1})$ to denote the probability that player $i$ chooses $a_i^t \in A_i(h^{t-1})$. I use $a^t = (a_1^t, \ldots, a_n^t) \in \times_{i=1}^n A_i(h^{t-1}) \equiv A(h^{t-1})$ to denote the action profile at period $t$ and $a_{-i}^t$ to denote the action profile at period $t$ without player $i$. If $a^t$ is the action profile realized at period $t$, then $h^t = (h^{t-1}, a^t)$. Finally, each player $i$ has a payoff function (in von Neumann-Morgenstern utilities) $u_i : \Theta \times \mathcal{H}^T \to \mathbb{R}$, and we let $u = (u_1, \ldots, u_n)$ be the profile of payoff functions. A multi-stage game of incomplete information, $\Gamma$, is defined by the tuple $\Gamma = \langle N, \Theta, \mathcal{H}, F, u \rangle$.

## 2.2 Dynamic Cognitive Hierarchy Solution

In the framework of the DCH solution, each player $i$ is endowed with a level of sophistication $\tau_i \in \mathbb{N}_0$ which is independently drawn from the distribution $P_i(\tau_i)$. Without loss of generality, I assume $P_i(\tau_i) > 0$ for all $i \in N$ and $\tau_i \in \mathbb{N}_0$. Let $\tau = (\tau_1, \ldots, \tau_n)$ represent the level profile and $\tau_{-i}$ denote the level profile without player $i$. Due to the independence, the level profile is drawn from a distribution $P : \mathbb{N}_0^{|N|} \to (0, 1)$ such that $P(\tau) = \Pi_{i=1}^n P_i(\tau_i)$. Following Lin and Palfrey (2024), I assume that the distributions $F$ and $P$ are independent, meaning that players' payoff-relevant types and levels of sophistication are drawn independently.

Each player $i$ holds a prior belief about their opponents' levels, which satisfies the property of *truncated rational expectations*. Specifically, for each $i$, $j \neq i$, and $k$, let $\hat{P}_{ij}^k(\tau_j)$ represent level-$k$ player $i$'s prior belief about player $j$'s level, satisfing

$$\hat{P}_{ij}^k(\tau_j) = \begin{cases} \frac{P_j(\tau_j)}{\sum_{m=0}^{k-1} P_j(m)} & \text{if } \tau_j < k \\ 0 & \text{if } \tau_j \geq k. \end{cases} \tag{1}$$

The intuition behind this specification is that, although level-$k$ players mistakenly believe that all others are at most level-$(k-1)$,[9] they still hold correct beliefs about the relative proportions of players who are less sophisticated than themselves.

In the DCH solution, a strategy profile refers to a level-dependent behavioral strategy profile of all levels of players. Let $\sigma_i^k$ be level-$k$ player $i$'s behavioral strategy, where level-0 players uniformly randomize at every information set.[10] That is, for every $i \in N$, $\theta_i \in \Theta_i$,

[9] The cognitive hierarchy specification aligns with behavioral and psychological evidence of overconfidence across various domains (see, for instance, Camerer and Lovallo, 1999; Moore and Healy, 2008; Enke et al., 2023).

[10] Uniform randomization is not the only way to model level-0 players' behavior; however, one compelling justifica-

$h \in \mathcal{H}\backslash\mathcal{H}^T$, and for any $a \in A_i(h)$, $\sigma_i^0(a|\theta_i, h) = 1/|A_i(h)|$.

At every history $h^t$, every level-$k \geq 1$ player $i$ forms a *joint belief* about all other players' types and levels.[11] Their posterior beliefs at history $h^t$ depend on the level-dependent strategy profile and the prior beliefs. To formalize the belief updating process, let $\sigma_j^{<k} = (\sigma_j^0, \ldots, \sigma_j^{k-1})$ be the profile of behavioral strategies adopted by the levels below $k$ of player $j$. Furthermore, let $\sigma_{-i}^{<k} = (\sigma_1^{<k}, \ldots, \sigma_{i-1}^{<k}, \sigma_{i+1}^{<k}, \ldots, \sigma_n^{<k})$ be the profile of behavioral strategies of the levels below $k$ of all players other than player $i$. Note that all strategic players believe every history is possible because $\hat{P}_{ij}^k(0) > 0$ for any $i, j \in N$ and $k > 0$, and $\sigma_j^0(a|\theta_j, h) > 0$ for any $j$, $\theta_j$, $h$ and $a \in A_j(h)$. Consequently, Bayes' rule can be applied to derive every level of players' posterior belief about other players' types and levels at every history.

Specifically, for any $i \in N$, $k \geq 1$ and $\theta_i \in \Theta_i$, a level-dependent strategy profile will induce the posterior belief $\mu^k(\theta_{-i}, \tau_{-i}|\theta_i, h)$ at every $h \in \mathcal{H}\backslash\mathcal{H}^T$ with $\mu^k(\theta_{-i}|\theta_i, h^{t-1})$ and $\mu^k(\tau_{-i}|\theta_i, h^{t-1})$ being level-$k$ player $i$'s marginal beliefs of other players' types and levels at history $h^{t-1}$, respectively. Lastly, for any $j \neq i$, let $\mu_j^k(\theta_j, \tau_j|\theta_i, h^{t-1})$ denote level-$k$ player $i$'s marginal belief about player $j$'s type and level at history $h^{t-1}$.

In the DCH solution, players correctly anticipate how they will update their posterior beliefs at all future histories of the game, i.e., players are strategic learners. Therefore, for any $i$, $k$, $\theta_i$ and any level-dependent strategy profile of others $\sigma_{-i}^{<k}$, type $\theta_i$ level-$k$ player $i$ believes the probability of $a_{-i}^t \in A_{-i}(h^{t-1})$ being chosen is

$$\tilde{\sigma}_{-i}^{<k}(a_{-i}^t|\theta_i, h^{t-1}) \equiv \sum_{\theta_{-i}\in\Theta_{-i}} \sum_{\{\tau_{-i}:\tau_j<k\ \forall j\neq i\}} \mu^k(\theta_{-i}, \tau_{-i}|\theta_i, h^{t-1}) \prod_{j\neq i} \sigma_j^{\tau_j}(a_j^t|\theta_j, h^{t-1}).$$

Furthermore, for every level of players, given lower-level players' strategies, they can compute the probability of any outcome being realized at any non-terminal history. In particular, for any $i \in N$, $\tau_i > 0$, $\theta \in \Theta$, $\sigma$, and $\tau_{-i}$ such that $\tau_j < \tau_i$ for any $j \neq i$, let $P_i^{\tau_i}(h^T|\theta, h^t, \tau_{-i}, \sigma_{-i}^{<\tau_i}, \sigma_i^{\tau_i})$ be level-$\tau_i$ player $i$'s belief about the conditional realization probability of $h^T \in \mathcal{H}^T$ at history $h^t \in \mathcal{H}\backslash\mathcal{H}^T$ if the type profile is $\theta$, the level profile is $\tau$, and player $i$ uses $\sigma_i^{\tau_i}$. Finally, level-$\tau_i$ player $i$'s expected payoff at any $h^t \in \mathcal{H}\backslash\mathcal{H}^T$ is:

$$\mathbb{E}u_i^{\tau_i}(\sigma|\theta_i, h^t) = \sum_{h^T\in\mathcal{H}^T} \sum_{\theta_{-i}\in\Theta_{-i}} \sum_{\{\tau_{-i}:\tau_j<\tau_i\ \forall j\neq i\}} \mu^{\tau_i}(\theta_{-i}, \tau_{-i}|\theta_i, h^t) P_i^{\tau_i}(h^T|\theta, h^t, \tau_{-i}, \sigma_{-i}^{<\tau_i}, \sigma_i^{\tau_i}) u_i(\theta_i, \theta_{-i}, h^T). \quad (2)$$

tion for its use is its universal applicability to all games in the same manner. In fact, the DCH solution is well-defined as long as level-0 players' behavioral strategy is fully mixed at every information set. See Section 7.4 in Lin and Palfrey (2024) for a detailed discussion.

[11] Level-1 players always believe other players are level-0 whose actions are uninformative about their types. Therefore, they don't update their beliefs about the levels and types of others.

The *DCH solution* of the game is defined as the level-dependent assessment $(\sigma^*, \mu^*)$, where $\sigma_i^{k*}(\cdot|\theta_i, h^t)$ maximizes (2) for all $i$, $k$, $\theta_i$, and $h^t \in \mathcal{H}\backslash\mathcal{H}^T$. The *DCH belief system* $\mu^*$ is induced by $\sigma^*$ via Bayes' rule. Additionally, it is assumed that players uniformly randomize over optimal actions when they are indifferent—a standard assumption in level-$k$ models to ensure the uniqueness of the solution.

**Remark 1.** *For one-stage games, the DCH solution reduces to the standard CH solution because one-stage games are essentially static games.*

## 2.3 General Properties of DCH in Multi-Stage Games of Incomplete Information

In this section, I first consider the case where the payoff-relevant types are independently drawn, i.e., $F(\theta) = \prod_{i\in N} F_i(\theta_i)$. Proposition 1 shows that for any multi-stage game of incomplete information, the DCH belief system is a product measure at every information set. That is, the posterior beliefs always remain independent across players.

**Proposition 1.** *For any multi-stage game of incomplete information $\Gamma$, any $h \in \mathcal{H}\backslash\mathcal{H}^T$, any $i \in N$, $\theta_i \in \Theta_i$, and for any $k \in \mathbb{N}$, if the prior distribution of types is independent across players, i.e., $F(\theta) = \prod_{i=1}^n F_i(\theta_i)$, then level-$k$ player $i$'s posterior belief about other players' types and levels at $h$ is independent across players. That is,*

$$\mu^k(\theta_{-i}, \tau_{-i}|\theta_i, h) = \prod_{j\neq i} \mu_j^k(\theta_j, \tau_j|\theta_i, h).$$

Proposition 1 extends the independence property shown by Lin and Palfrey (2024) from games of perfect information to multi-stage games of incomplete information. This generalization relies on that (1) the actions are perfectly observed and (2) players are able to perform Bayesian inferences.

In multi-stage games of incomplete information, where actions are publicly revealed at the end of each period, Proposition 1 demonstrates that DCH strategic players understand that an individual player's action does not reveal any information about any other player's private type or level of sophistication. However, as noted by Lin and Palfrey (2024), this property does not hold for general extensive games of imperfect information. When actions are not perfectly observed, the marginal beliefs about other players' levels may become correlated (see Section 7.2 of Lin and Palfrey 2024).

Besides, the ability to perform Bayesian inferences plays a crucial role in maintaining the independence property. In other behavioral solution concepts, such as the Cursed Sequential

Equilibrium proposed by Fong et al. (2025), where players fail to understand the correlation between other players' actions and their private information, they may mistakenly believe that others' actions are informative about another player's private information, even though the actions are perfectly observed and the prior distribution of types is independent across players.

Next, I consider the case where the prior distribution of types is not independent across players. When types are correlated, a player's actions convey information not only about their own private type but also about the private types of other players whose types are correlated with theirs. To address correlated types, the original game can be transformed into another game (with independent types) through a specific transformation.

For any multi-stage game of incomplete information $\Gamma$, following Myerson (1985) and Fudenberg and Tirole (1991), we can consider a corresponding transformed game $\hat{\Gamma}$ where the prior distribution of types is the product of independent uniform marginal distributions. That is, for any $\theta \in \Theta$, $\hat{F}(\theta) = 1/\prod_{i=1}^{n} |\Theta_i|$. Additionally, the utility functions are transformed as follows: $\hat{u}_i(\theta_i, \theta_{-i}, h^T) = F(\theta_{-i}|\theta_i)u_i(\theta_i, \theta_{-i}, h^T)$.

Proposition 2 shows that the DCH level-dependent behavioral strategy profile is invariant under the transformation between the transformed and original game, suggesting that the independence assumption of the types is without loss of generality.

**Proposition 2.** *The level-dependent assessment $(\hat{\sigma}, \hat{\mu})$ is the DCH solution of the transformed game (with independent types) if and only if the level-dependent assessment $(\sigma, \mu)$ is the DCH solution of the original game (with correlated types) where $\sigma = \hat{\sigma}$ and for any $i \in N$, $\theta_i \in \Theta_i$, $k > 0$, and $h^t \in \mathcal{H}\backslash\mathcal{H}^T$,*

$$\mu^k(\theta_{-i}, \tau_{-i}|\theta_i, h^t) = \frac{F(\theta_{-i}|\theta_i)\hat{\mu}^k(\theta_{-i}, \tau_{-i}|\theta_i, h^t)}{\sum_{\theta'_{-i}} \sum_{\{\tau'_{-i}:\tau'_j<k \ \forall j\neq i\}} F(\theta'_{-i}|\theta_i)\hat{\mu}^k(\theta'_{-i}, \tau'_{-i}|\theta_i, h^t)}.$$

Given that the DCH solution is solved recursively from the bottom of the hierarchy, it is natural to conjecture that achieving this result would require a family of level-dependent transformations. Therefore, it is surprising that Proposition 2 is derived using a single transformation for all levels of players. The intuition behind this result lies in the fact that, since types and levels are independently determined, players cannot infer others' types based on their knowledge of others' levels. As a result, the transformation remains level-independent.

Moreover, Proposition 2 shows that at any information set $I_i = (\theta_i, h^t)$, player $i$'s belief about a specific type-level profile $(\theta_{-i}, \tau_{-i})$ is proportional to the prior belief of $\theta_{-i}$ conditional on $\theta_i$. In addition, if $F(\theta_{-i}|\theta_i) \to 1$, i.e., $\theta_i$ is almost perfectly correlated with $\theta_{-i}$, then

$$\mu^k(\theta_{-i}, \tau_{-i}|\theta_i, h^t) = \frac{F(\theta_{-i}|\theta_i)\hat{\mu}^k(\theta_{-i}, \tau_{-i}|\theta_i, h^t)}{\sum_{\theta'_{-i}} \sum_{\{\tau'_{-i}:\tau'_j<k\ \forall j\neq i\}} F(\theta'_{-i}|\theta_i)\hat{\mu}^k(\theta'_{-i}, \tau'_{-i}|\theta_i, h^t)}$$
$$\to \frac{\hat{\mu}^k(\theta_{-i}, \tau_{-i}|\theta_i, h^t)}{\sum_{\{\tau'_{-i}:\tau'_j<k\ \forall j\neq i\}} \hat{\mu}^k(\theta_{-i}, \tau'_{-i}|\theta_i, h^t)} = \hat{\mu}^k(\theta_{-i}, \tau_{-i}|\theta_i, h^t),$$

implying that the belief in the transformed game aligns with the belief in the original game. Intuitively speaking, if the types are almost perfectly correlated, then the remaining information to be learned is solely others' levels. Since the DCH behavioral strategy profile is invariant under the transformation, the belief about others' levels will also be invariant under the transformation.

Finally, since level-0 players totally mix at every history, strategic players can never rule out the possibility that all other players are level-0. Consequently, in the DCH solution, every non-terminal history remains on-path, and the marginal belief of each level of player about others' types always remains full support. Furthermore, because it is always possible for other players to be level-0, common knowledge of rationality is never achieved in DCH, regardless of how sophisticated the players are. This distinction underscores the fundamental differences between DCH and equilibrium theory as solution concepts.

In the next section, I will illustrate another important property of DCH, the violation of invariance under strategic equivalence, using the dirty-faces game. In DCH, since level-0 players uniformly randomize at every information set, their behavioral strategies might not be outcome-equivalent if the cardinality of action sets changes. Consequently, this non-equivalence triggers a chain reaction affecting all higher-level players, as DCH is solved recursively from the bottom of the hierarchy.[12]

# 3 DCH Analysis of the Dirty-Faces Game

The dirty-faces game was originally introduced as a mathematical puzzle to illustrate iterative rationality by Littlewood (1953). Since then, various versions of the game have been used as illustrative examples in numerous theoretical studies on common knowledge (see, for example, Binmore and Brandeburger, 1988; Fudenberg and Tirole, 1993; Geanakoplos, 1994; Liu, 2008). In this paper, I focus on the simplest variant: the two-person dirty-faces game.

[12] According to Thompson et al. (1952) and Elmes and Reny (1994), two extensive games share the same reduced normal form if and only if they can be transformed into each other using a small set of elementary transformations. Specifically, Elmes and Reny (1994) propose three such transformations: INT, COA, and ADD, which preserve perfect recall. Because DCH is sensitive to the cardinality of action sets, it varies under COA while remaining invariant under INT and ADD. See Battigalli (2024) for a detailed discussion.

There are two players $N = \{1, 2\}$ and there are up to $2 \leq T < \infty$ periods. At the beginning of the game, each player $i$ is randomly assigned a face type, denoted as $x_i$, which can be either $x_i = O$ (representing a clean face) or $x_i = X$ (representing a dirty face). The face types are i.i.d. drawn from the distribution $d = \Pr(x_i = X) = 1 - \Pr(x_i = O)$ where $d > 0$ represents the probability of having a dirty face. After the face types are determined, each player $i$ can observe the other player's face type $x_{-i}$ but not their own face. Hence, player $i$'s private information is the other player's face type $x_{-i}$. Furthermore, if at least one player has a dirty face, a public announcement is made, informing both players of this fact.

If there is no announcement, it is common knowledge to both players that both faces are clean. To avoid triviality, I focus on the case where an announcement is made.

After seeing the other player's face type and the announcement, in each period, every player $i$ simultaneously chooses to "Wait" ($W$) or "Claim" (to have a dirty face, $C$) and their actions are revealed at the end of each period. The game will end after any period where some player chooses $C$ or after period $T$. The last period of the game is called the "terminal period," and both players' payoffs are determined by their own face types and their actions in the terminal period.

Suppose period $t$ is the terminal period. If player $i$ chooses $W$ in the terminal period, his payoff for this game is 0 regardless of his face type. On the other hand, if player $i$ chooses $C$ to terminate the game, he will receive $\alpha > 0$ if his face is dirty and $-1$ if his face is clean. Besides, payoffs are discounted with a common discount factor $\delta \in (0, 1)$ per period. That is, if player $i$ claims in period $t$ and $x_i = X$, player $i$ will receive $\delta^{t-1}\alpha$, but player $i$ will receive $-\delta^{t-1}$ if $x_i = O$. To make players unattractive to gamble if they do not have additional information except for the prior, following Weber (2001) and Bayer and Chan (2007), I assume that

$$d\alpha - (1 - d) < 0 \iff 0 < \bar{\alpha} \equiv \frac{d\alpha}{1 - d} < 1, \tag{3}$$

which guarantees that it is strictly dominated to choose $C$ in period 1 when seeing a dirty face. Thus, a two-person dirty-faces game is defined by a tuple $\langle T, \delta, \bar{\alpha}c6f, d\rangle$ where $(\delta, \bar{\alpha}) \in (0, 1)^2$.

With common knowledge of rationality, the unique equilibrium can be solved through the following iterative reasoning: When player $i$ sees a clean face, the public announcement will lead him to realize that his own face is dirty and claim in period 1. On the other hand, when player $i$ sees a dirty face, he will wait in period 1 because of the uncertainty about his own face. However, if player $-i$ also waits in period 1, player $i$ will then recognize that his own face is dirty and claim in period 2, as player $i$ knows that if his own face were clean, player $-i$ would have claimed in period 1.

When implementing this game in a laboratory experiment, the natural approach is to specify this game as a *sequential* dirty-faces game and allow subjects to make decisions period-by-period, following the rules described above, using the direct-response method. Alternatively, the other approach is the strategy method which specifies this game as a *simultaneous* dirty-faces game—after seeing the other's face and the announcement, players simultaneously decide a "plan" which specifies the period to claim or always wait. From the standard game-theoretic perspective, the sequential and simultaneous dirty-faces game are strategically equivalent as they share the same reduced normal form. In the following, I will demonstrate that the DCH solution varies in these two versions of the game, illustrating the violation of invariance under strategic equivalence of DCH.

## 3.1 DCH Solution for the Sequential Dirty-Faces Games

In the sequential dirty-faces game, since there are $T$ periods, a behavioral strategy for player $i$ is a mapping from the observed face type ($x_{-i} \in \{O, X\}$) and the period to the probability of choosing $C$, denoted by $\sigma_i : \{O, X\} \times \{1, \ldots, T\} \to [0, 1]$.

For the sake of simplicity, I assume that each player $i$'s level is i.i.d. drawn from the distribution $p = (p_k)_{k=0}^{\infty}$ where $p_k > 0$ for all $k$. In the DCH solution, each player's optimal behavioral strategy is level-dependent. Let the behavioral strategy of level-$k$ player $i$ be $\sigma_i^k$. Following previous notations, let $\mu^k(x_i, \tau_{-i}|x_{-i}, t)$ be level-$k$ player $i$'s belief about their own face and the level of the other player, conditional on observing $x_{-i}$ and being at period $t$. Level-0 players will uniformly randomize everywhere, so $\sigma_i^0(x_{-i}, t) = 1/2$ for all $x_{-i}$ and $t$.

Proposition 3 fully characterizes the DCH solution for the sequential dirty-faces games. When observing a clean face, a strategic player can immediately figure out that his face is dirty. Therefore, DCH coincides with the equilibrium prediction when $x_{-i} = O$. However, if a player sees a dirty face and the other player waits in period 1, he cannot tell his face type for sure, no matter how sophisticated he is. Instead, he will believe that he is more likely to have a dirty face as the game continues. As a result, conditional on observing a dirty face, level-$k \geq 2$ players will claim as long as the reward $\bar{\alpha}$ is high enough or the discount rate $\delta$ is sufficiently low. Otherwise, they will wait for more evidence.

**Proposition 3.** *For any sequential two-person dirty-faces game, the level-dependent strategy profile of the DCH solution satisfies that for any $i \in N$,*

1. $\sigma_i^k(O, t) = 1$ *for any* $k \geq 1$ *and* $1 \leq t \leq T$.

2. $\sigma_i^1(X, t) = 0$ *for any* $1 \leq t \leq T$. *Moreover, for any* $k \geq 2$,

   *(1)* $\sigma_i^k(X, 1) = 0$,

*(2) for any* $2 \leq t \leq T-1$, $\sigma_i^k(X,t)=1$ *if and only if*

$$\bar{\alpha} \;\geq\; \frac{\left[\left(\frac{1}{2}\right)^{t-1} - \left(\frac{1}{2}\right)^{t}\delta\right] p_0}{\left[\left(\frac{1}{2}\right)^{t-1} - \left(\frac{1}{2}\right)^{t}\delta\right] p_0 + (1-\delta)\sum_{j=1}^{k-1} p_j},$$

*(3)* $\sigma_i^k(X,T)=1$ *if and only if*

$$\bar{\alpha} \;\geq\; \frac{\left(\frac{1}{2}\right)^{T-1} p_0}{\left(\frac{1}{2}\right)^{T-1} p_0 + \sum_{j=1}^{k-1} p_j}.$$

To gain insights into the model, I analyze the behavior of level-1 and 2 players. Level-1 players believe the other player is non-strategic, implying that the other player's actions do not provide any information about their face type. In other words, for level-1 players, the announcement and their own observations are the only relevant source of information. Thus, level-1 players in each period will behave exactly the same as in period 1—they will claim when seeing a clean face, and wait when seeing a dirty face.

Level-2 players will also claim to have a dirty face in period 1 upon seeing a clean face, and will wait in period 1 (which is a strictly dominant strategy) if they see a dirty face. However, due to the presence of level-0 players, even if the game proceeds to period 2, level-2 players remain uncertain about their own face type. Specifically, in period 2, level-2 players form a joint belief about their own face type $x_i$ and the other player's level $\tau_{-i}$, denoted by $\mu^2(x_i, \tau_{-i}|X,2)$, where

$$\mu^2(X,0|X,2) = \frac{\left(\frac{1}{2}\right) dp_0}{\left(\frac{1}{2}\right) p_0 + dp_1}, \qquad \mu^2(O,0|X,2) = \frac{\left(\frac{1}{2}\right)(1-d)p_0}{\left(\frac{1}{2}\right) p_0 + dp_1},$$
$$\mu^2(X,1|X,2) = \frac{dp_1}{\left(\frac{1}{2}\right) p_0 + dp_1}, \qquad \mu^2(O,1|X,2) = 0.$$

In other words, as the game progresses beyond period 2, level-2 players will realize that it is *impossible* for the other player to be level-1 while their own face is clean. Therefore, for any period $2 \leq t \leq T$, the marginal belief of level-2 players about having a dirty face is

$$\mu^2(X|X,t) = \underbrace{\frac{\left(\frac{1}{2}\right)^{t-1} dp_0}{\left(\frac{1}{2}\right)^{t-1} p_0 + dp_1}}_{=\mu^2(X,0|X,t)} + \underbrace{\frac{dp_1}{\left(\frac{1}{2}\right)^{t-1} p_0 + dp_1}}_{=\mu^2(X,1|X,t)} = \frac{d\left[\left(\frac{1}{2}\right)^{t-1} p_0 + p_1\right]}{\left(\frac{1}{2}\right)^{t-1} p_0 + dp_1},$$

which is increasing in $t$, suggesting that level-2 players are more certain about having a dirty face in later periods. This is level-2 players' benefit of waiting when seeing a dirty face. However, the cost of waiting is that the other player may randomly end the game (if the other is level-0) and the payoff is discounted. Therefore, level-2 players' tradeoff is analogous to the sequential sampling problem of Wald (1947)—they decide the *optimal stopping period* to claim. The optimal stopping period depends on the parameters $\bar{\alpha}$ and $\delta$, as well as the distribution of levels. This is in sharp contrast with the equilibrium prediction that the equilibrium prediction is independent of the parameters.

Particularly, for any period $2 \leq t \leq T$, level-2 player $i$'s expected payoff of claiming to have a dirty face is $\mathbb{E}u_i^2(C|t) := \delta^{t-1}\left[\alpha\mu^2(X|X,t) - \mu^2(O|X,t)\right]$. At period $T$, the last period of the game, it is optimal to claim if and only if

$$\mathbb{E}u_i^2(C|T) \geq 0 \iff \alpha \geq \frac{\mu^2(O|X,T)}{\mu^2(X|X,T)} \iff \bar{\alpha} \geq \frac{\left(\frac{1}{2}\right)^{T-1} p_0}{\left(\frac{1}{2}\right)^{T-1} p_0 + p_1}.$$

For any other period $2 \leq t' \leq T-1$, it is optimal to claim at period $t'$ only if $\mathbb{E}u_i^2(C|t') \geq \Pr(t'+1|X,t')\mathbb{E}u_i^2(C|t'+1)$, where $\Pr(t'+1|X,t')$ is level-2 player $i$'s belief about the probability that player $-i$ would wait in period $t'$.[13] Rearranging the inequality yields the condition stated in Proposition 3. Furthermore, the proof in the Supplemental Appendix shows that these conditions are not only necessary but also sufficient to pin down level-2 players' optimal stopping periods. By induction on the levels, it can be shown that no matter how sophisticated the players are, the behavior is characterized by the solution of a sequential sampling problem.

Proposition 3 characterizes the level-dependent behavioral strategies. Alternatively, the DCH solution can be characterized by the level-dependent stopping period (given observing $x_{-i}$), which is formally defined in Definition 1.

**Definition 1** (Stopping Period). *For any sequential two-person dirty-faces game and its DCH level-dependent strategy profile $\sigma$, let $\hat{\sigma}_i^k(x_{-i})$ be level-$k$ player $i$'s earliest period to claim to have a dirty face conditional on observing $x_{-i}$ for any $k \geq 1$ and $i \in N$. Specifically,*

$$\hat{\sigma}_i^k(x_{-i}) = \begin{cases} \min\left\{t' : \sigma_i^k(x_{-i}, t') = 1\right\}, & \textit{if } \exists\, t \textit{ s.t. } \sigma_i^k(x_{-i}, t) = 1 \\ T+1, & \textit{otherwise.} \end{cases}$$

[13] $\Pr(t'+1|X,t')$ is the probability that the other player chooses to wait in period $t'$, which is

$$\Pr(t'+1|X,t') = 0.5 \cdot \mu^2(0|X,t') + 1 \cdot \mu^2(1|X,t') = \frac{\left(\frac{1}{2}\right)^{t'} p_0 + dp_1}{\left(\frac{1}{2}\right)^{t'-1} p_0 + dp_1}.$$

With Definition 1, Corollary 1 is a direct consequence of Proposition 3. If $x_{-i} = O$, every strategic level of players will know their face is dirty and claim to have a dirty face in period 1, viz. $\hat{\sigma}_i^k(O) = 1$ for every $k \geq 1$. In contrast, if $x_{-i} = X$, Corollary 1 shows that the optimal stopping period is monotonically decreasing in $k$, implying that higher-level players tend to claim sooner.

**Corollary 1.** *For any sequential two-person dirty-faces game, the DCH level-dependent strategy profile $\sigma$ can be equivalently characterized by level-dependent stopping periods. Moreover, for any $i \in N$, we know*

1. $\hat{\sigma}_i^k(O) = 1$ *for any* $k \geq 1$,
2. $\hat{\sigma}_i^1(X) = T + 1$, *and* $\hat{\sigma}_i^k(X) \geq 2$ *for all* $k \geq 2$.
3. $\hat{\sigma}_i^k(X)$ *is weakly decreasing in* $k$.

To summarize, I illustrate the DCH optimal stopping periods of level-2 and level-infinity players when seeing a dirty face, i.e., $\hat{\sigma}_i^2(X)$ and $\hat{\sigma}_i^\infty(X)$. Because the set of dirty-faces games is described by $(\delta, \bar{\alpha})$, it is simply the unit square on the $(\delta, \bar{\alpha})$-plane. For the illustrative purpose, I consider $T = 5$ and the distribution of levels follows Poisson(1.5), which is an empirically regular prior according to Camerer et al. (2004).

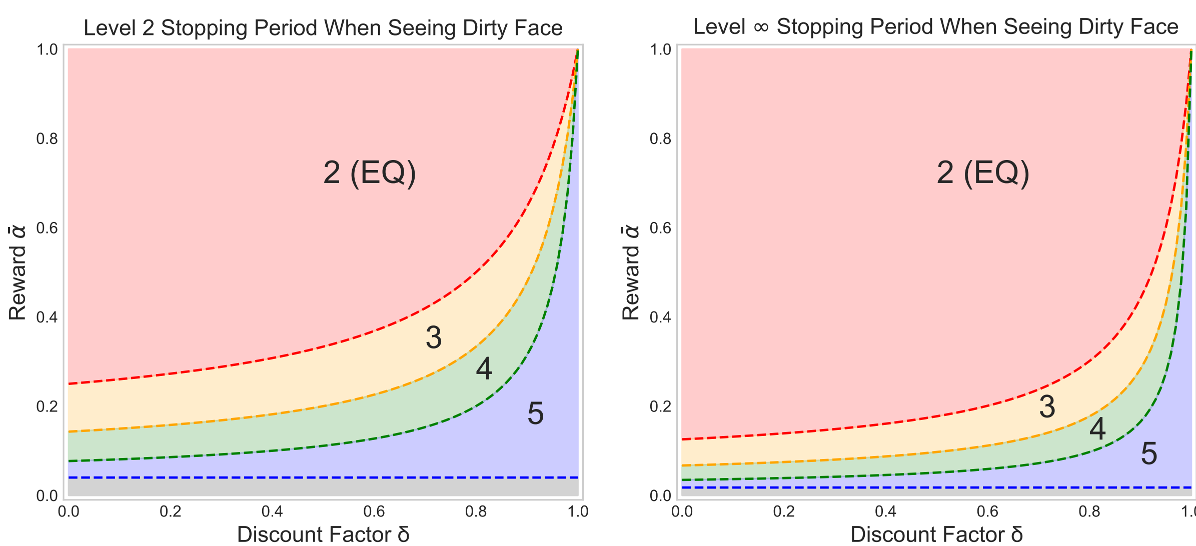



Figure 1: DCH stopping periods in sequential dirty-faces games for level-2 (Left) and level-$\infty$ players (Right) as $x_{-i} = X$ where $T = 5$ and the distribution of levels follows Poisson(1.5).

The DCH stopping periods can be solved according to Proposition 3 and are plotted in Figure 1, showing that DCH predicts that when seeing a dirty face, it is possible for strategic players to choose any stopping period in $\{2, 3, 4, 5, 6\}$, depending on the parameters $\bar{\alpha}$ and $\delta$. For instance, level-2 players will claim in period 2 (red area) if and only if $\bar{\alpha} \geq \frac{2-\delta}{8-7\delta}$.

Lastly, DCH predicts the comparative statics that the optimal stopping period is weakly decreasing in $\bar{\alpha}$ and weakly increasing in $\delta$ for any level-$k \geq 2$. The intuition is that when $\bar{\alpha}$ is larger or $\delta$ is smaller, waiting becomes more costly, which causes the players to claim earlier with a less certain belief about their own face type.

## 3.2 DCH Solution for the Simultaneous Dirty-Faces Games

In contrast, the strategically equivalent simultaneous dirty-faces game is essentially a one-period game where players simultaneously choose an action from the set $S = \{1, \dots, T+1\}$. Action $t \leq T$ represents the plan to wait from period 1 to $t-1$ and claim in period $t$. Action $T+1$ is the plan to always wait. In the simultaneous dirty-faces game, a mixed strategy for player $i$ is a mapping from the observed face type ($x_{-i} \in \{O, X\}$) to a probability distribution over the action set. The mixed strategy is denoted by $\tilde{\sigma}_i : \{O, X\} \to \Delta(S)$. Suppose $(s_i, s_{-i})$ is the action profile. If $s_i \leq s_{-i}$, then the payoff for player $i$ is computed as the case where player $i$ claims in period $s_i$; if $s_i > s_{-i}$, then player $i$'s payoff is 0.

To illustrate how the DCH solution varies between the two versions of the game, I again assume that each player's level is i.i.d. drawn from the distribution $p = (p_k)_{k=0}^{\infty}$ where $p_k > 0$ for all $k$. Level-0 players will uniformly randomize, regardless of what they observe, so $\tilde{\sigma}_i^0(x_{-i}) = \frac{1}{T+1}$ for all $i, x_{-i}$. Since level-$k \geq 1$ players will generically choose pure strategies, I slightly abuse the notation to use $\tilde{\sigma}_i^k(x_{-i})$ to denote the pure strategies.[14]

Proposition 4 is parallel to Proposition 3 that characterizes the DCH solution for the simultaneous dirty-faces games. The intuition is similar to the analysis of sequential dirty-faces games. When observing a clean face, players can figure out their face types immediately. Hence, they will choose the strictly dominant strategy $\tilde{\sigma}_i^k(O) = 1$ for all $k \geq 1$. On the other hand, when observing a dirty face, players have to make *hypothetical* inferences about their face types and the other player's level of sophistication.

**Proposition 4.** *For any simultaneous two-person dirty-faces game, the level-dependent strategy profile of the DCH solution satisfies that for any $i \in N$,*

1. $\tilde{\sigma}_i^k(O) = 1$ *for any* $k \geq 1$.

2. $\tilde{\sigma}_i^1(X) = T + 1$. *Moreover, for any* $k \geq 2$,

   *(1)* $\tilde{\sigma}_i^k(X) \geq 2$,

[14] That is, for any $t \in \{1, 2, \dots, T, T+1\}$, I use $\tilde{\sigma}_i^k(x_{-i}) = t$ to denote the degenerated distribution: $\tilde{\sigma}_i^k(x_{-i})(t) = 1$ and $\tilde{\sigma}_i^k(x_{-i})(t') = 0 \ \forall\, t' \neq t$.

*(2) for any $2 \leq t \leq T-1$, $\tilde{\sigma}_i^k(X) \leq t$ if and only if*

$$\bar{\alpha} \;\geq\; \frac{\left[\frac{T+2-t}{T+1} - \frac{T+1-t}{T+1}\delta\right] p_0}{\left[\frac{T+2-t}{T+1} - \frac{T+1-t}{T+1}\delta\right] p_0 + (1-\delta)\sum_{j=1}^{k-1} p_j},$$

*(3) $\tilde{\sigma}_i^k(X) \leq T$ if and only if*

$$\bar{\alpha} \;\geq\; \frac{\frac{2}{T+1}p_0}{\frac{2}{T+1}p_0 + \sum_{j=1}^{k-1} p_j}.$$

The characterization is similar to Proposition 3. When observing a dirty face, strategic players will not choose 1, i.e., they will wait in period 1. Instead, they will claim in period $t$ if and only if the reward $\bar{\alpha}$ is sufficiently high or the discount rate $\delta$ is sufficiently low. However, the critical threshold for $\bar{\alpha}$ differs, indicating a violation of invariance under strategic equivalence. Although DCH makes different quantitative predictions in the sequential and simultaneous dirty-faces games, it makes a similar qualitative prediction that higher-level players tend to claim earlier than lower-level players. This is proven in Corollary 2.

**Corollary 2.** *For any simultaneous two-person dirty-faces game, the DCH level-dependent strategy profile $\tilde{\sigma}$ satisfies that for any $i \in N$ and any $k \geq 2$, $\tilde{\sigma}_i^k(X)$ weakly decreases in $k$.*

## 3.3 The Violation of Invariance under Strategic Equivalence

As discussed in the previous sections, DCH predicts that players may behave differently in two strategically equivalent dirty-faces games. This result arises from the fact that when the cardinality of the action sets differs, the behavioral strategies of level-0 players may no longer be outcome-equivalent. This discrepancy triggers a chain reaction that influences the behavior of higher-level players, since the DCH solution is solved recursively. In this subsection, I characterize *how* changes in the cardinality of the action sets influence behavior in dirty-faces games.

In the sequential game, the cardinality of each player $i$'s action set (conditional on each $x_{-i}$) is $2^T$, whereas in the simultaneous game, it is $T+1$. This difference in cardinalities of action sets leads to distinct behavior among level-0 players. For instance, in the first period, level-0 players will claim to have a dirty face with probability $1/2$ in the sequential game, but only with probability $1/(T+1)$ in the simultaneous game. While this discrepancy does not affect level-1 players, it significantly influences how level-2 (and more sophisticated) players update their beliefs about their own face types.

To characterize this effect, I partition the set of dirty-faces games based on the stopping rules of each level in the sequential and simultaneous versions, respectively. First, for any level-$k \geq 1$, let $\mathcal{E}_t^k$ denote the set of sequential games in which level-$k$ players will claim no later than period $t$ when they observe a dirty face. In other words, $(\delta, \bar{\alpha}) \in \mathcal{E}_t^k$ if and only if $\hat{\sigma}_i^k(X) \leq t$.[15] Second, the set of dirty-faces games can alternatively be partitioned by the stopping rules of each level in the simultaneous games, i.e., $\tilde{\sigma}_i^k(X)$. For any $t \geq 1$ and $k \geq 1$, let $\mathcal{S}_t^k$ be the set of dirty-faces games such that $\tilde{\sigma}_i^k(X) \leq t$. Using this notation, Proposition 5 characterizes how DCH differs between the sequential and simultaneous games through set inclusions between $\mathcal{E}_t^k$ and $\mathcal{S}_t^k$.

**Proposition 5.** *Consider any $T \geq 2$ and the set of two-person dirty-faces games. For any level-$k \geq 2$, the following relationships hold.*

1. $\mathcal{S}_T^k \subset \mathcal{E}_T^k$.
2. $\mathcal{S}_t^k \subset \mathcal{E}_t^k$ *for any* $[\ln(T+1)/\ln 2] \leq t \leq T-1$.
3. *There is no set inclusion relationship between* $\mathcal{S}_t^k$ *and* $\mathcal{E}_t^k$ *for* $2 \leq t < [\ln(T+1)/\ln 2]$. *Moreover, for any $i \in N$, there exists $\overline{\delta}(T,t) \in (0,1)$ such that $t = \hat{\sigma}_i^k(X) \leq \tilde{\sigma}_i^k(X)$ if $\delta \leq \overline{\delta}(T,t)$ and $\hat{\sigma}_i^k(X) \geq \tilde{\sigma}_i^k(X) = t$ if $\delta > \overline{\delta}(T,t)$. Specifically,*

$$\overline{\delta}(T,t) = \frac{(2^t-2)(T+1)-(t-1)2^t}{(2^t-1)(T+1)-t2^t}.$$

Proposition 5 formally compares the DCH solutions for the sequential and simultaneous dirty-faces games. First, the inclusion $\mathcal{S}_T^k \subset \mathcal{E}_T^k$ for any $k \geq 2$ implies that, upon observing a dirty face, players are more likely to claim before the game ends in the sequential game than in the simultaneous game. Moreover, conditional on reaching any period $t \geq \ln(T+1)/\ln 2$, DCH predicts that players tend to claim earlier in the sequential game than in the simultaneous game, *regardless* of the payoff parameters or the level distribution. However, this does not mean that players always claim earlier in the sequential game. The third result shows that when the horizon is sufficiently long and players are sufficiently patient—specifically, when $\delta > \overline{\delta}(T,t)$—it is possible for players to claim *later* in the sequential game. More surprisingly, the threshold $\overline{\delta}(T,t)$ is independent of the level of sophistication and the prior distribution over levels, suggesting that the structural difference between the two versions of the game has a uniform impact across all levels of players.

[15] By Corollary 1, since level-1 players never claim upon seeing a dirty face, we have $\mathcal{E}_t^1 = \emptyset$ for all $t = 1, \ldots, T$, and $\mathcal{E}_{T+1}^1 = (0,1)^2$. For higher-level players, the partitions depend on the payoff parameters and the distribution of levels. For example, when the distribution of levels follows Poisson(1.5), the set $\mathcal{E}_2^2$ is characterized by $(\delta, \bar{\alpha}) \in \mathcal{E}_2^2 \iff (2-\delta)/(8-7\delta) \leq \bar{\alpha} < 1$ where $0 < \delta < 1$.

To illustrate this proposition, consider the case where $T = 5$ and the distribution of levels follows Poisson(1.5). Figure 2 displays the partitions for level-2 players in both the sequential and simultaneous games. By Proposition 5, we find that $\mathcal{S}_t^k \subset \mathcal{E}_t^k$ for any $k \geq 2$ and $3 \leq t \leq 5$, indicating that DCH predicts players—upon seeing a dirty face and with the game continuing to period 3 or later—are more likely to claim earlier in the sequential game than in the simultaneous game. However, no set inclusion holds between $\mathcal{S}_2^k$ and $\mathcal{E}_2^k$, as illustrated in the right panel of Figure 2, where I overlay the boundaries of $\mathcal{S}_2^k$ and $\mathcal{E}_2^k$.

Numerically, by Propositions 3 and 4, the sets $\mathcal{E}_2^2$ and $\mathcal{S}_2^2$ can be characterized by:

$$(\delta, \bar{\alpha}) \in \mathcal{E}_2^2 \iff \bar{\alpha} \geq \frac{\left(\frac{1}{2} - \frac{1}{4}\delta\right) e^{-1.5}}{\left(\frac{1}{2} - \frac{1}{4}\delta\right) e^{-1.5} + (1-\delta)1.5e^{-1.5}} = \frac{2-\delta}{8-7\delta},$$
$$(\delta, \bar{\alpha}) \in \mathcal{S}_2^2 \iff \bar{\alpha} \geq \frac{\left(\frac{5}{6} - \frac{2}{3}\delta\right) e^{-1.5}}{\left(\frac{5}{6} - \frac{2}{3}\delta\right) e^{-1.5} + (1-\delta)1.5e^{-1.5}} = \frac{5-4\delta}{14-13\delta}.$$

These boundaries intersect at $\delta = 0.8$, suggesting that when $\delta < 0.8$, level-2 players are more likely to claim earlier in the sequential game; when $\delta > 0.8$, the reverse holds. In fact, for any level-$k \geq 2$, Proposition 5 guarantees that the boundaries of $\mathcal{E}_2^k$ and $\mathcal{S}_2^k$ also intersect at $\delta = 0.8$, indicating that the cutoff $\bar{\delta}(5, 2)$ is the same across all levels.

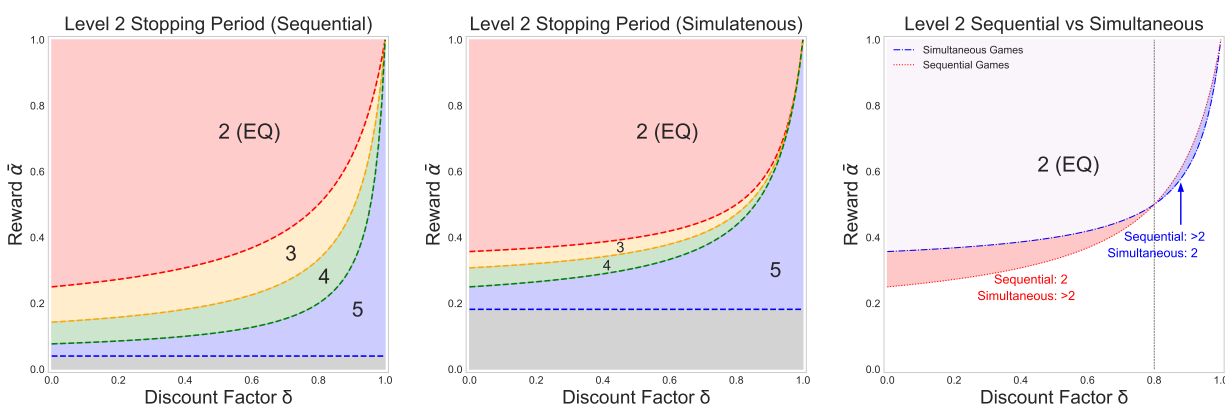



Figure 2: (Left) DCH stopping periods in sequential dirty-faces games for level-2. (Middle) DCH stopping periods in simultaneous dirty-faces games for level-2. (Right) Overlay of the sets $\mathcal{E}_2^2$ and $\mathcal{S}_2^2$. The figures are plotted under the assumption that the distribution of levels follows Poisson(1.5).

Lastly, as the maximum horizon $T$ increases, the cardinalities of the action sets in the sequential and simultaneous games become increasingly distinct. Consequently, the behavior of level-0 players diverges even further between the sequential and simultaneous games. As $T \to \infty$, Proposition 5 implies for any period $t \geq 2$ and level-$k \geq 2$, $\mathcal{S}_t^k$ and $\mathcal{E}_t^k$ do not have set inclusion relationship, suggesting that higher-level players do not definitely learn their face types earlier in one game or another. Their behavior depends on the parameters $(\delta, \bar{\alpha})$. The result is formally presented in Corollary 3.

**Corollary 3.** *When $T \to \infty$, for any $t \geq 2$ and $k \geq 2$, there is no set inclusion relationship between $\mathcal{S}_t^k$ and $\mathcal{E}_t^k$. Specifically, if $\delta < \overline{\delta}^*(t)$, then $t = \hat{\sigma}_i^k(X) \leq \tilde{\sigma}_i^k(X)$; and if $\delta > \overline{\delta}^*(t)$, then $\hat{\sigma}_i^k(X) \geq \tilde{\sigma}_i^k(X) = t$ where $\overline{\delta}^*(t) = [2^t - 2]/[2^t - 1]$.*

In summary, this section characterizes the DCH solutions for the sequential and simultaneous two-person dirty-faces games. I show that because level-0 players uniformly randomize across all actions, their behavior is not outcome-equivalent in the two versions of the game, leading level-$k \geq 2$ players to behave differently. Importantly, the assumption of uniform randomization is not critical to this result—as long as level-0 players' non-strategic behavior is not outcome-equivalent, the violation of invariance under strategic equivalence of DCH will still occur.

**Remark 2.** *When there are more than two players, DCH predicts a larger difference between the two versions, as the boundaries between the sequential and simultaneous games become further apart. In the working paper version of this paper (Lin, 2023), I characterize the DCH solutions for three-player three-period games and find that level-$k \geq 3$ players tend to learn their face types earlier in the sequential game.*

# 4 Experimental Design, Hypotheses and Procedures

As demonstrated in the previous section, DCH makes various predictions about how people's behavior would vary with the timing (sequential vs. simultaneous) and the payoff structures of the dirty-faces games. To test these predictions, I conduct a laboratory experiment on two-person dirty-faces games tailored to evaluate the DCH solution.

Specifically, the primary goal of the experiment is to measure the violation of invariance under strategic equivalence and understand how it interacts with the payoff structures. Furthermore, the variation of the payoff structures provides the opportunity to explore the sensitivity of behavior to payoffs in both the sequential and simultaneous versions of the game. Lastly, the stylized facts found in this experiment will help identify the strengths and the weaknesses of the DCH solution and alternative theories.

The theoretical analysis of DCH suggests that the main challenge in designing the experiment lies in the fact that the magnitude of the difference between the two versions depends on the payoff structure and the distribution of levels, which remains unknown before the experiment is run. To address this, I first estimate the distribution of levels using the dirty-faces game experimental data collected by Bayer and Chan (2007), and then choose the game parameters to maximize the diagnosticity based on the calibration results.

## 4.1 Calibration

The dirty-faces game experiment by Bayer and Chan (2007) is implemented under the direct response method with two treatments: two-person two-period games and three-person three-period games. In both treatments, the prior probability of having a dirty face is 2/3, the discount factor $\delta$ is 4/5, and the reward $\alpha$ is 1/4.[16] I will focus on the data from two-person games because this environment is the closest to my experiment.[17] A detailed analysis of the data can be found in Appendix A.

There are 42 subjects (from two sessions) in the two-person treatment of Bayer and Chan (2007). At the beginning of the experiment, the computer randomly matches two subjects into a group. Subjects play 14 rounds of dirty-faces games against the same opponent, with the face types in each round being independently drawn according to the prior probabilities. In each round, an announcement is made on the screen to both subjects if there is at least one person having a dirty face (type $X$). At the end of each round, subjects are told their own payoffs from that round and they are paid with the sum of the earnings of all 14 rounds.

In the calibration exercise, I exclude the data from the situation where there is no public announcement[18], resulting in 690 observations at the information set level. Following previous notations, I use $(t, x_{-i})$ to denote the situation where subject $i$ sees type $x_{-i}$ at period $t$. Table 1 reports the empirical frequency of choosing *claim* at each information set, revealing that the behavior is inconsistent with the prediction of standard equilibrium theory, particularly when observing a dirty face.

Following the literature on the cognitive hierarchy theory, I assume the prior distribution of levels follows a Poisson distribution. In the Poisson-DCH model, each individual $i$'s level is identically and independently drawn from $(p_k)_{k=0}^{\infty}$ where $p_k = e^{-\tau}\tau^k/k!$ for all $k \in \mathbb{N}_0$ and $\tau > 0$. Once the distribution of levels is specified, DCH makes a precise prediction about the aggregate choice frequency at each information set.[19] The rationale for estimating the

[16] In Bayer and Chan (2007), the payoff of correctly claiming a dirty face is 100 ECU (experimental currency unit) and the penalty of wrongly claiming a dirty face is $-400$ ECU. Therefore, the relative reward of correctly claiming a dirty face $\alpha = 1/4$ can be obtained by normalizing the payoffs.

[17] Weber (2001)'s dataset consists of two experiments where experiment 2 is comparable with Bayer and Chan (2007)'s design. However, there are much fewer observations in this experiment than Bayer and Chan (2007) and there is no discount factor, making this dataset less ideal for the purpose of calibration.

[18] If there is no public announcement, it is common knowledge that both subjects' faces are clean.

[19] The aggregate choice frequency can be constructed as follows. Consider any game, any player $i$, any information set $\mathcal{I}_i$, and any available action $c_i$ at this information set. Let $P_k(c_i|\mathcal{I}_i)$ represent the probability of level $k$ player $i$

Poisson-DCH model is to find $\tau$, estimated using the maximum likelihood method, which minimizes the difference between the choice frequencies predicted by DCH and the empirical frequencies. See Appendix A.2 for the details on the construction of the likelihood function.

It is worth remarking that since $\tau$ is the mean (and variance) of the Poisson distribution, the economic interpretation of $\tau$ is as the average level of sophistication among the population. Additionally, another property of the Poisson-DCH model is that as $\tau \to \infty$, the aggregate choice frequencies predicted by DCH converge to the equilibrium predictions. This provides a second interpretation of $\tau$: the higher the value of $\tau$, the closer the predictions are to the equilibrium. See Proposition O.1 in the Supplemental Appendix for the proof.

Table 1: Estimation Results for Two-Person Dirty-Faces Games

| | $(t, x_{-i})$ | $N$ | $\sigma_i^*(t, x_{-i})$ | $\hat{\sigma}_i(t, x_{-i})$ | DCH | Standard CH |
|---|---|---|---|---|---|---|
| $\sigma_i(t, x_{-i})$ | $(1, O)$ | 123 | 1.000 | 0.943 | 0.859 | 0.791 |
| | $(2, O)$ | 6 | 1.000 | 0.500 | 0.500 | 0.500 |
| | $(1, X)$ | 391 | 0.000 | 0.210 | 0.141 | 0.104 |
| | $(2, X)$ | 170 | 1.000 | 0.618 | 0.503 | 0.477 |
| Parameter | $\tau$ | | | | **1.269** | 1.161 |
| | S.E. | | | | **(0.090)** | (0.095) |
| Fitness | LL | | | | -360.75 | -381.46 |
| | AIC | | | | 723.50 | 764.91 |
| | BIC | | | | 728.04 | 769.45 |
| Vuong Test | | | | | | 6.517 |
| p-value | | | | | | $< 0.001$ |

Note: The equilibrium and the empirical frequencies of $C$ at each information set are denoted as $\sigma_i^*$ and $\hat{\sigma}_i$, respectively. There are 294 games (rounds × groups).

Table 1 reports the estimation results of the Poisson-DCH model. Additionally, I estimate the standard Poisson-CH model by Camerer et al. (2004) as a benchmark.[20] Comparing the fitness of these models, I find that the log-likelihood of DCH is significantly higher than standard CH (Vuong test p-value $< 0.001$), suggesting that DCH outperforms the standard CH in capturing the empirical pattern. Besides, the estimated $\tau$ of Poisson-DCH falls within the range of commonly observed $\tau$ in various environments, with a value of 1.269.

choosing $c_i$ at $\mathcal{I}_i$. Additionally, let $f(k|\mathcal{I}_i, \tau)$ be the posterior distribution of levels at information set $\mathcal{I}_i$. The choice frequency predicted by DCH for action $c_i$ at information set $\mathcal{I}_i$ is the aggregation of choice probabilities from all levels, weighted by the proportion $f(k|\mathcal{I}_i, \tau)$:

$$\mathcal{D}(c_i|\mathcal{I}i, \tau) \equiv \sum_{k=0}^{\infty} f(k|\mathcal{I}_i, \tau) P_k(c_i|\mathcal{I}_i, \tau).$$

[20]The logit-AQRE proposed by McKelvey and Palfrey (1998) is also estimated. The likelihood scores between Poisson-DCH and logit-AQRE are not significantly different. See Appendix A for the details.

In the following, I will design the experiment by treating Poisson(1.269) as the true prior distribution of levels.

## 4.2 Games and Hypotheses

In this experiment, I employ a between-subject design where each participant is assigned to either the "sequential treatment" (using the direct-response method) or the "simultaneous treatment" (using the strategy method). To observe potential heterogeneity in stopping periods, I set the maximum length to be $T = 5$ for both treatments.

Assessing whether the difference between the two treatments is challenging because level 1 players—the most common types of players according to the calibration result—will behave the same under the two treatments. When observing a dirty face, they will always wait in both the sequential and simultaneous games. Therefore, to diagnose the predictivity of DCH, the game parameters are chosen to make level 2 players behave differently under different representations. The behavioral change of level 2 players (around 22.6% based on the calibration) is anticipated to yield a sizable treatment effect.

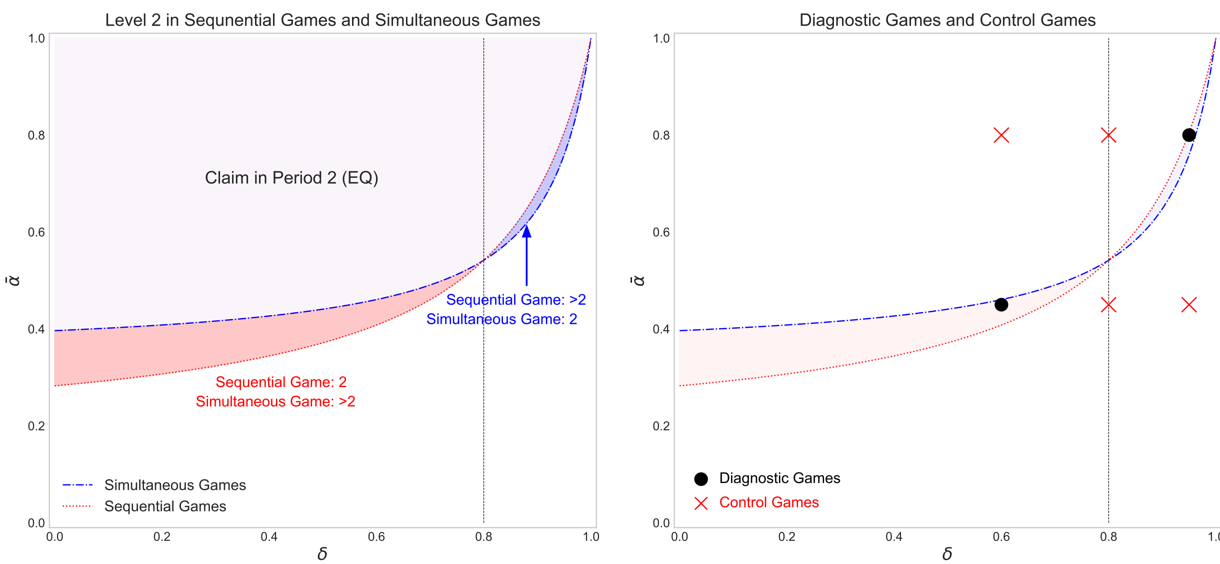



Figure 3: (Left) The set of dirty-faces games where at information set $(2, X)$, level 2 players behave differently in the two versions when $T = 5$ and the distribution of levels follows Poisson(1.269). (Right) The two diagnostic games and the four control games in the experiment.

According to Proposition 5, DCH predicts the existence of a set of dirty-faces games in which level 2 players exhibit different behavior in the sequential and simultaneous games at information set $(2, X)$. The left panel of Figure 3 illustrates this set of games when the distribution of levels follows Poisson(1.269). From this figure, we can observe the following:

(1) For $\delta < 0.8$, there is a range of games (red area) where level 2 players choose to claim at $(2, X)$ in the sequential games but not in the simultaneous games.

(2) For $\delta = 0.8$, level 2 (and more sophisticated) players behave the same in the sequential and simultaneous games.

(3) For $\delta > 0.8$, there is a range of games (blue area) where level 2 players choose to claim at $(2, X)$ in the simultaneous games but not in the sequential games.

Guided by DCH, I consider the following six dirty-faces games $(\delta, \bar{\alpha})$ as depicted in the right panel of Figure 3.

The set of games consists of two *diagnostic games* where $(\delta, \bar{\alpha}) = (0.6, 0.45)$ and $(0.95, 0.8)$ and four *control games* where $(\delta, \bar{\alpha}) = (0.6, 0.8), (0.8, 0.45), (0.8, 0.8)$ and $(0.95, 0.45)$. DCH predicts in the diagnostic games, level 2 players will behave differently in two treatments, but not in the control games. This variation allows us to examine the interplay between the violation of invariance under strategic equivalence varies with the payoff structures.

DCH makes several predictions about the comparative statics. First, by Proposition 3 and 4, DCH predicts that no matter in sequential games or in simultaneous games, when observing a dirty face, players will choose to claim *earlier* when $\delta$ is smaller or $\bar{\alpha}$ is higher. An implication is that in both treatments, at information set $(2, X)$, players are more likely to claim when $\delta$ decreases or $\bar{\alpha}$ increases.

**Hypothesis 1.** *In both the sequential and simultaneous treatments, at information set* $(2, X)$*, the empirical frequency of choosing* $C$ *is higher when* $\delta$ *decreases or* $\bar{\alpha}$ *increases.*

Besides, DCH makes a specific prediction regarding the relative magnitude of the treatment effect among these six games. First, in the DCH solution, part of the treatment effect is attributed to the difference in level 0 players' strategies between the two treatments. In the sequential games, level 0 players uniformly randomize at every information set, resulting in a conditional probability to claim at $(2, X)$ is $1/2$. Yet, in the simultaneous games, level 0 players uniformly randomize across all reduced contingent strategies, leading to a conditional probability to claim at $(2, X)$ is $1/5$. In other words, the difference in level 0 players' strategies generates a mechanical effect that increases the likelihood of players choosing to claim at $(2, X)$ in the sequential games. Because in all four control games, strategic players behave the same at $(2, X)$ under two representations, DCH predicts the magnitude of the violation of invariance under strategic equivalence will be similar in the control games. Particularly, in the game $(\delta, \bar{\alpha})$, the treatment effect can be quantified by computing the difference between the conditional probabilities of choosing to claim at $(2, X)$ in the sequential version and the simultaneous version. This difference is denoted by $\Delta(\delta, \bar{\alpha})$.[21]

Second, in the game where $(\delta, \bar{\alpha}) = (0.6, 0.45)$, level 2 players will claim at $(2, X)$ in the sequential version but not in the simultaneous version. As a result, DCH predicts that the difference between the two treatments in this diagnostic game will be stronger compared to the effect observed in the control games. On the contrary, in the game where $(\delta, \bar{\alpha}) = (0.95, 0.8)$, level 2 players will claim at $(2, X)$ in the simultaneous version but not in the sequential version. This offsets the mechanical effect caused by level 0 players. The expected differences based on the calibration results are summarized below.

**Hypothesis 2.** *Based on the calibration results, the expected differences are:*

$$\underset{\substack{\parallel \\ 31.15\%}}{\Delta(0.6, 0.45)} > \underset{\substack{\parallel \\ 7.4\%}}{\Delta(0.6, 0.8)} = \underset{\substack{\parallel \\ 7.4\%}}{\Delta(0.8, 0.8)} \approx \underset{\substack{\parallel \\ 4.82\%}}{\Delta(0.8, 0.45)} \approx \underset{\substack{\parallel \\ 3.26\%}}{\Delta(0.9, 0.45)} > \underset{\substack{\parallel \\ -18.93\%}}{\Delta(0.95, 0.8)}.$$

### 4.3 Experimental Procedures

The experimental sessions were conducted at the Experimental Social Science Laboratory (ESSL) located on the campus of the University of California, Irvine. Subjects were recruited from the general undergraduate population, from all majors. Experiments were conducted through oTree software (Chen et al., 2016). I conducted 10 sessions with a total of 118 subjects. No subject participated in more than one session. Each session lasted around 45 minutes, and the average earnings was \$33.36, including the \$10 show-up fee (max \$52 and min \$10).

Subjects were given instructions at the beginning and the instructions were read aloud. Subjects were allowed to ask any questions during the whole instruction process. The questions were answered so that every one can hear. Afterwards, they had to answer several comprehension questions on the computer screen in order to proceed. The instructions for both the sequential treatment and the simultaneous treatments are identical except for the instructions about the choices and the feedback after each game. The instructions for both treatments can be found in the Supplemental Appendix.

[21] In the sequential version, the observed conditional probability of claiming at $(2, X)$ is simply the empirical $\sigma_i(2, X)$. For the simultaneous version, the conditional probability can be computed from the empirical $\tilde{\sigma}_i(X)$ by $\tilde{\sigma}_i(2, X) \equiv \Pr(\tilde{\sigma}_i(X) = 2)/\sum_{t=2}^{6} \Pr(\tilde{\sigma}_i(X) = t)$. Therefore, the treatment effect is quantified by the (empirical) difference between $\sigma_i(2, X)$ and $\tilde{\sigma}_i(2, X)$, i.e., $\Delta \equiv \sigma_i(2, X) - \tilde{\sigma}_i(2, X)$.

Table 2: List of Game Parameters Implemented in the Experiment

| | Game Parameters | | | | Normalized Probabilities | |
|---|---|---|---|---|---|---|
| | $\delta$ | $\alpha$ | $p$ | $\bar{\alpha}$ | one X, one O | two X |
| Diagnostic Game 1 | 0.60 | 0.225 | 0.67 | 0.45 | 0.25 | 0.50 |
| Diagnostic Game 1' | 0.60 | 0.150 | 0.75 | 0.45 | 0.20 | 0.60 |
| Diagnostic Game 2 | 0.95 | 0.400 | 0.67 | 0.80 | 0.25 | 0.50 |
| Diagnostic Game 2' | 0.95 | 0.267 | 0.75 | 0.80 | 0.20 | 0.60 |
| Control Game 1 | 0.60 | 0.400 | 0.67 | 0.80 | 0.25 | 0.50 |
| Control Game 1' | 0.60 | 0.267 | 0.75 | 0.80 | 0.20 | 0.60 |
| Control Game 2 | 0.80 | 0.225 | 0.67 | 0.45 | 0.25 | 0.50 |
| Control Game 2' | 0.80 | 0.150 | 0.75 | 0.45 | 0.20 | 0.60 |
| Control Game 3 | 0.80 | 0.400 | 0.67 | 0.80 | 0.25 | 0.50 |
| Control Game 3' | 0.80 | 0.267 | 0.75 | 0.80 | 0.20 | 0.60 |
| Control Game 4 | 0.95 | 0.225 | 0.67 | 0.45 | 0.25 | 0.50 |
| Control Game 4' | 0.95 | 0.150 | 0.75 | 0.45 | 0.20 | 0.60 |

Each session comprised 12 games with different $(\delta, \alpha, p)$ configurations, as summarized in Table 2.[22] The sequence of these games was randomized, and in each game, subjects were randomly paired into groups. The draws for player types were independent, and the protocol was common knowledge. To prevent any framing effect, the "dirty face" and the "clean face" were labelled as "red" and "white" in the instruction, respectively. Besides, the actions were labelled as "I'm red" and "wait." Finally, to avoid situations where both faces are clean, the probabilities were normalized to ensure that having two clean faces was impossible.[23]

After observing the other's face type, subjects were asked to simultaneously choose their actions. In the sequential treatment, subjects simultaneously chose either "I'm red" or "wait." If both subjects chose to wait, the game would proceed to the next period and they were asked to choose again. The game ended after the period where some one chose "I'm red" or after period 5. On the other hand, in the simultaneous treatment, subjects simultaneously chose one of the six plans (the period to choose "I'm red" or always wait) and the plans were implemented by the computer. At the end of each match, the subjects were informed of their own payoffs, the true types and the histories of the game.[24]

Lastly, subjects were paid in cash based on their total points earned from the 12 games.

[22]Notice that the parameters are selected such that each $(\delta, \bar{\alpha})$ is played twice.

[23]For example, in the game with $p = 2/3$, subjects were informed that the probability of one dirty face and one clean face was $1/4$, and the probability of two dirty faces was $1/2$. Therefore, if the other's face was clean, the subject could infer that his own face was dirty. Conversely, if the other's face was dirty, the subject's belief about his own face being dirty was $2/3$.

[24]To control for the amount of feedback in both treatments, in the simultaneous treatment, subjects would learn the other's exact plan if the other chose "I'm red" earlier or at the same period; otherwise, they would be told that the other subject was *later than you*.

The highest possible earnings of each game was 100 points.[25] The conversion rate was two US dollars for every 100 points. Following previous dirty face game experiments (Weber, 2001; Bayer and Chan, 2007), each subject was provided an endowment of 900 points at the beginning of the experiment to prevent early bankruptcy, and they would only receive the show-up fee if the total point is negative.

# 5 Experimental Results

## 5.1 Aggregate-Level Analysis

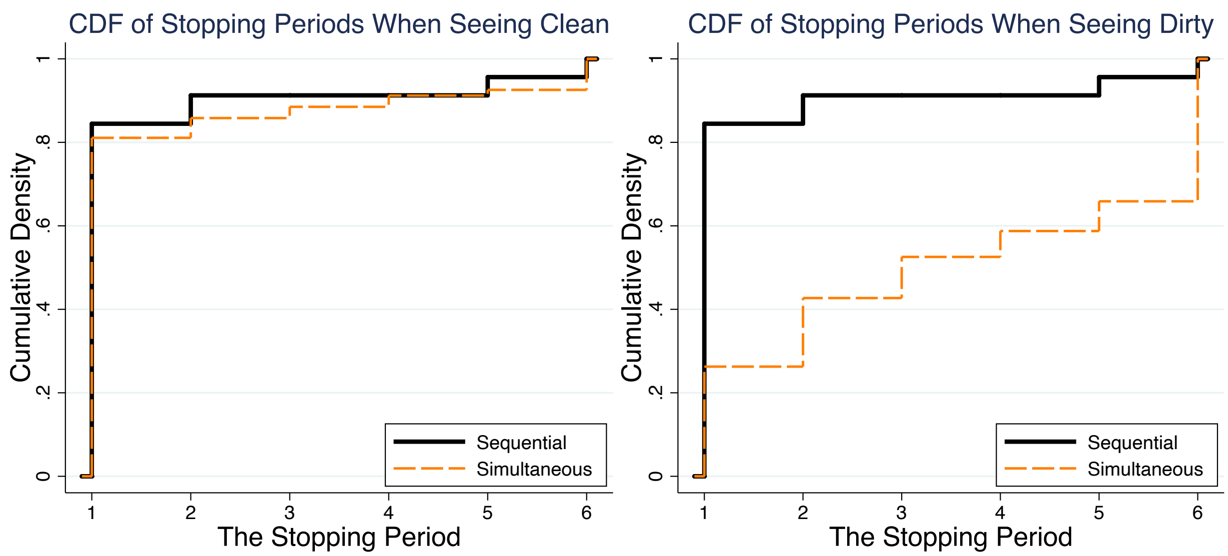



Figure 4: The CDFs of the stopping periods when players see either a clean face (left panel) or a dirty face (right panel). The black solid and the orange dashed lines are the distributions in the sequential and the simultaneous treatments, respectively.

The data includes two treatments (sequential and simultaneous) with 60 subjects in the sequential treatment and 58 subjects in the simultaneous treatment. Each subject participates in 12 games, resulting in 1024 observations for the sequential treatment and 1979 observations for the simultaneous treatment at the information set level.[26] Figure 4 provides a comprehensive overview of the data by plotting the distribution of stopping periods in both

[25]That is, a correct claim in the first period would yield 100 points, while an incorrect claim in the first period would result in a penalty of $100/\alpha$ points.

[26]For the simultaneous treatment, the choice data at the information set level are implied by the contingent strategies. For instance, choosing the contingent strategy "claim at period 4" implies that the subject will wait from period 1 to period 3 and claim in period 4.

treatments, aggregating across all payoff configurations. This analysis considers scenarios where players encounter either a clean face or a dirty face.[27]

A few key observations emerge from this figure. First, when players see a clean face, their behavior is consistent across both treatments. A majority of players seem to understand that their face is dirty and claim in period 1. Second, when players encounter a dirty face, it is evident that the distribution of stopping periods in the simultaneous treatment first-order stochastically dominates the distribution in the sequential treatment, implying that players are more inclined to claim earlier in the sequential treatment. Furthermore, a striking pattern in the right panel of Figure 4 is the prevalence of the "always wait" strategy in the simultaneous treatment, chosen by approximately 36.7% of participants. This is in stark contrast to the sequential treatment, where the proportion of participants employing the "always wait" strategy is only about 13.9%.

Focusing on data from the first two periods, Figure 5 presents the empirical frequencies of choosing $C$ at each information set. At information set $(O, 1)$, behavior does not differ significantly across treatments (Ranksum test p-value = 0.4423). In both treatments, the frequency of choosing $C$ exceeds 80%, indicating that the majority of subjects understand that choosing $C$ in the first period, upon seeing a clean face, is a strictly dominant strategy.

Moreover, the behavior at information set $(O, 2)$ provides valuable insights into the rationale behind DCH. From the perspective of DCH, the information set $(O, 2)$ is reached only when a player is level-0. Therefore, DCH predicts that the frequencies of $C$ in the sequential and simultaneous treatments should be 50% and 20%, respectively. As depicted in Figure 5, the frequencies of $C$ in the sequential and simultaneous treatments are 43.8% and 25%, respectively,[28] aligning with the prediction of DCH. Despite the limited number of observations at information set $(O, 2)$, this result provides supportive evidence for rationale of DCH.

On the other hand, when players see a dirty face, they need to make inferences about their own faces either from the opponent's actions or hypothetically. However, regardless of the treatment, claiming in period 1 is strictly dominated. Comparing the empirical frequencies of choosing $C$ at information set $(X, 1)$, we find that players in the simultaneous treatment are less likely to choose $C$ (Ranksum test p-value = 0.0641). This observation is consistent with DCH, as level 0 players in the simultaneous treatment are less likely to claim at information

[27]In the sequential treatment, the cumulative density of stopping periods is derived from the choice probability at each information set. For example, the probability of stopping in period 1 corresponds to the empirical frequency of choosing $C$. Similarly, the probability of stopping in period 2 is the product of the empirical frequency of choosing $W$ in period 1 and the empirical frequency of choosing $C$ in period 2. The probabilities for other stopping periods are calculated in a similar manner.

[28]A similar pattern is also found in Bayer and Chan (2007). In their dataset, the frequency of choosing $C$ at information set $(O, 2)$ is exactly 50%, which coincides with the prediction of DCH.

set $(X, 1)$. Furthermore, a significant treatment effect is detected at period 2. The frequency of choosing $C$ at information set $(X, 2)$ in the sequential treatment is 60.0%, while the frequency in the simultaneous treatment is 22.3% (Ranksum test p-value $< 0.0001$).

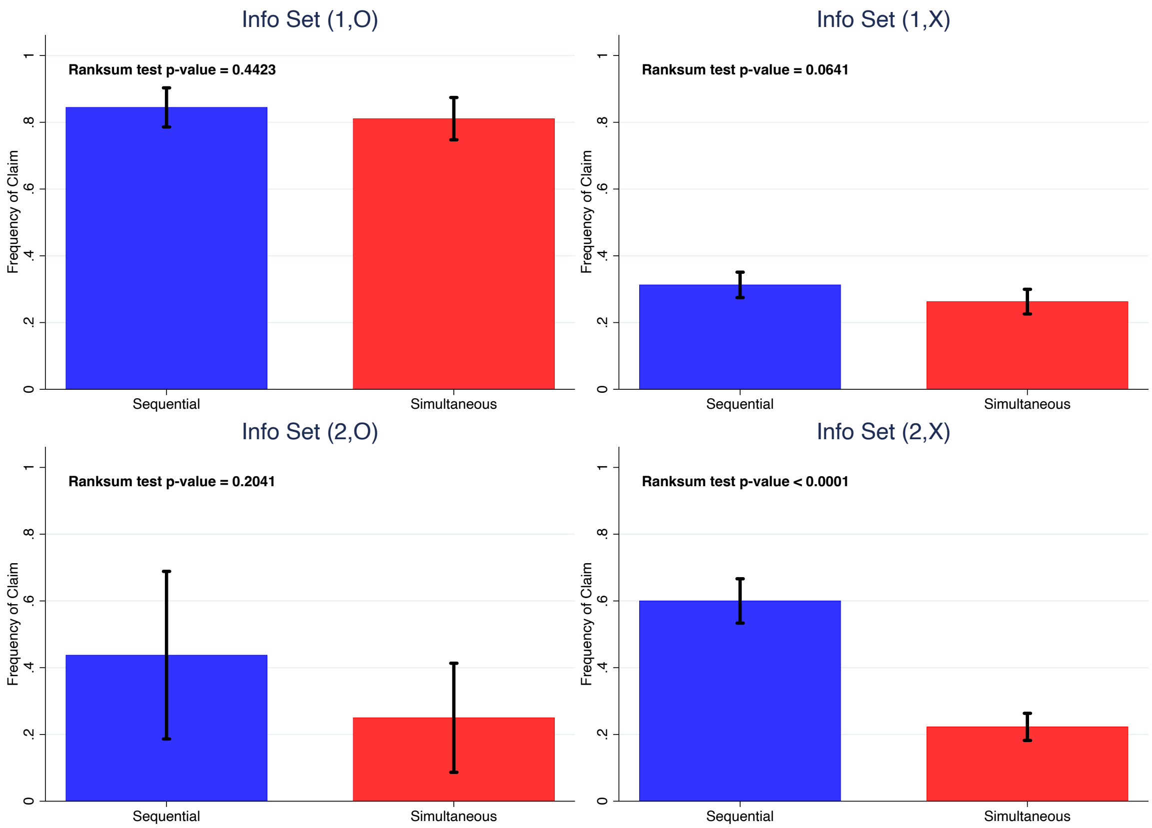



Figure 5: The empirical frequencies of $C$ and 95% CI at each information set in period 1 and 2, aggregating across all configurations. Each panel represents an information set. The blue bars are the frequencies of the sequential treatment and the red bars are the frequencies of the simultaneous treatment.

**Result 1.** *(1) When observing a clean face in both treatments, over 80% of the subjects choose $C$ in period 1, the strictly dominant strategy. Additionally, the behavior at information set $(O, 2)$ aligns with the prediction of DCH about level-0 players' behavior. (2) When players observe a dirty face, a significant difference emerges: they are more likely to claim at information set $(X, 2)$ in the sequential treatment. Furthermore, the most prevalent strategy in the simultaneous treatment when players see a dirty face is to "always wait."*

The supplementary analysis can be found in Appendix B.1. In the following, I will focus on information set $(2, X)$, where a strong treatment effect is found, and I will test two hypotheses related to the sensitivity of behavior to the payoff structures and the interplay between the payoff structures and the magnitude of the effect.

## 5.2 The Payoff Effect

Focusing on information set $(2, X)$, DCH predicts that in both treatments, players' behavior is sensitive to the payoff configurations. Specifically, DCH predicts that the empirical frequencies of choosing $C$ at information set $(2, X)$ will exhibit a monotonic relationship with $\delta$ and $\bar{\alpha}$. To test this prediction, I perform Kruskal-Wallis ranksum tests on the sequential and simultaneous treatments separately.

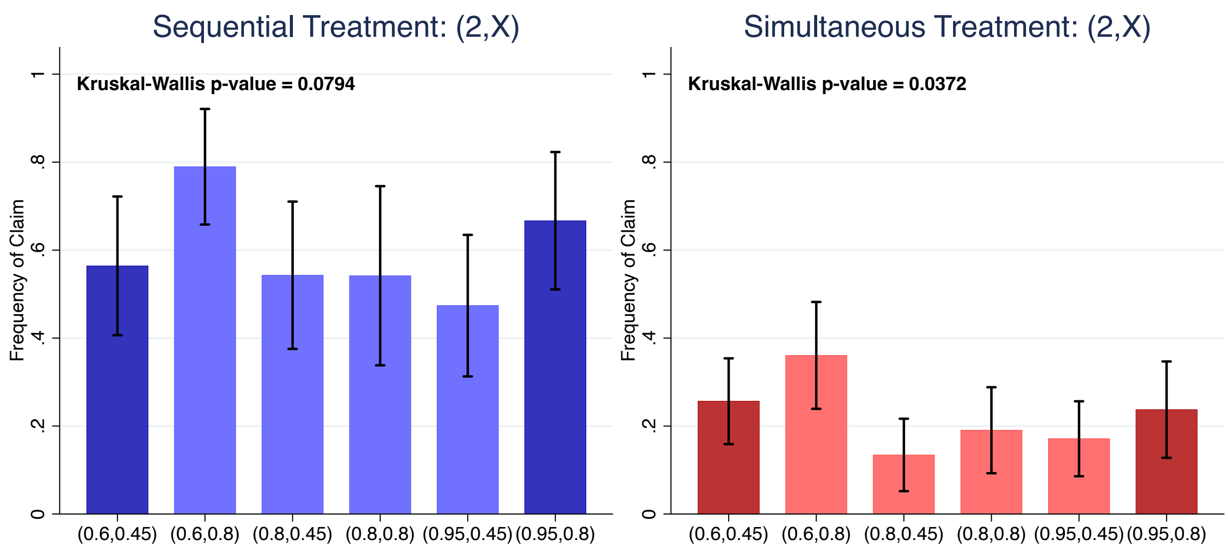



Figure 6: The empirical frequencies of $C$ and 95% CI at information set $(2, X)$ in each payoff configuration $(\delta, \bar{\alpha})$. The data from the sequential and the simultaneous treatments are plotted in the left and the right panel, respectively.

In the sequential treatment, we find that the null hypothesis is marginally rejected ($\chi^2(5) = 9.856$, p-value $= 0.0794$), suggesting that behavior is influenced by variations in payoff structures. Furthermore, we can observe from the left panel of Figure 6 that the frequency of choosing $C$ weakly increases with $\bar{\alpha}$ for any $\delta$. This monotonic pattern aligns with the prediction of DCH.

Similarly, the null hypothesis is rejected for the simultaneous treatment ($\chi^2(5) = 11.831$, p-value $= 0.0372$), indicating that behavior in the simultaneous treatment significantly varies with the payoff parameters. Once again, we can observe from Figure 6 that for each $\delta$, the frequency of choosing $C$ weakly increases with $\bar{\alpha}$, aligning with DCH.

**Result 2.** *The behavior at information set $(2, X)$ in both treatments significantly varies with payoffs, aligning with the qualitative predictions of DCH.*

## 5.3 The Violation of Invariance under Strategic Equivalence

The behavior in both treatments significantly varies with the payoff structures. Additionally, the difference in behavior between the two treatments also varies with the payoff structures. This variability allows us to examine the predictions of DCH.

First, the left panel of Figure 7 displays the joint distribution of the empirical frequencies of choosing $C$ at $(2, X)$ between the two treatments, where each point represents one payoff configuration. From the figure, we can observe that all six points are below the 45-degree line, implying that players are more likely to claim at $(2, X)$ in the sequential treatment than in the simultaneous treatment, regardless of the payoff configuration.

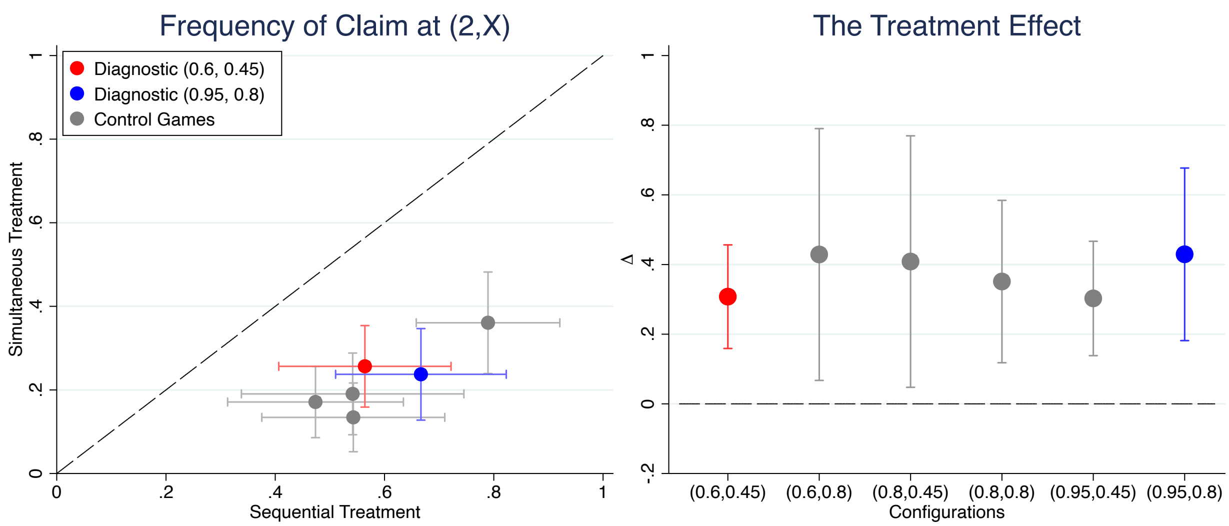



Figure 7: (Left) The empirical frequencies of $C$ and 95% CI at information set $(2, X)$ in each payoff configuration $(\delta, \bar{\alpha})$. (Right) The difference in the frequencies of $C$ at information set $(2, X)$ between the sequential and simultaneous treatments for each payoff configuration $(\delta, \bar{\alpha})$ with the 95% CIs. The standard errors are clustered at the session level.

Second, for each payoff configuration $(\delta, \bar{\alpha})$, I calculate $\Delta(\delta, \bar{\alpha})$, which represents the difference in the empirical frequencies of choosing $C$ at $(2, X)$ between both treatments. The results are shown in the right panel of Figure 7. Focusing on the diagnostic game where $(\delta, \bar{\alpha}) = (0.6, 0.45)$, we observe a significant treatment effect with a magnitude of $\Delta(0.6, 0.45) = 30.77\%$ (95% CI $= [15.90\%, 45.64\%]$, p-value $= 0.001$). This is highly consistent with the prediction of the calibrated DCH.

Furthermore, when focusing on the four control games, we find that the magnitudes of the treatment effects are similar in these games, which aligns with the qualitative predictions of DCH. However, these magnitudes are much stronger than the predictions of calibrated DCH, ranging from $\Delta(0.95, 0.45) = 30.26\%$ (95% CI = $[13.85\%, 46.68\%]$, p-value = 0.002) to $\Delta(0.6, 0.8) = 42.88\%$ (95% CI = $[6.73\%, 79.03\%]$, p-value = 0.025). Lastly, in the second diagnostic game with $(\delta, \bar{\alpha}) = (0.95, 0.8)$, DCH predicts a negative treatment effect based on the calibration results. However, the observed empirical difference for this game is $\Delta(0.95, 0.8) = 42.94\%$ (95% CI = $[18.17\%, 67.71\%]$ and p-value = 0.004), which is inconsistent with the quantitative prediction of the calibrated DCH.

**Result 3.** *In the diagnostic game with $(\delta, \bar{\alpha}) = (0.6, 0.45)$, the frequency of $C$ at information set $(2, X)$ is 30.77% higher in the sequential treatment compared to the simultaneous treatment. In the diagnostic game with $(\delta, \bar{\alpha}) = (0.95, 0.8)$, the difference is 42.94%. Furthermore, treatment effects are detected in all control games, with magnitudes exceeding the predictions of calibrated DCH.*

In summary, when analyzing the interplay between the violation of strategic equivalence and payoff structures, we observe that while calibrated DCH captures some qualitative patterns, the observed magnitudes are significantly larger. This suggests that the observed behavior might result from both the violation of mutual consistency and other behavioral biases. To delve deeper into this aspect, in the next subsection, I will compare DCH with other behavioral models that relax other requirements of the standard equilibrium theory.

## 5.4 Structural Estimation and Model Comparison

DCH relaxes the requirement of mutual consistency in sequential equilibrium while still adhering to the requirements of best response and Bayesian inference. Can the empirical pattern be better explained by relaxing other requirements? To assess the relaxation of two other requirements, I estimate the "Quantal Cursed Sequential Equilibrium (QCSE),"[29] which is a hybrid model combining AQRE and CSE, thereby relaxing the requirements of best response and Bayesian inference.

QCSE assumes that players are unable to fully understand how other players' actions depend on their private information.[30] In particular, for any strategy profile $\sigma$, any player $i$

[29] See Appendix B.2 for a detailed description of the model.

[30] In the context of the dirty-faces game, QCSE assumes that players do not fully recognize how the other player's actions depend on the observed face type.

and any information set $\mathcal{I}_i = (\theta_i, h^{t-1})$, the *average behavioral strategy of player* $-i$ is

$$\bar{\sigma}_{-i}\left(a_{-i}^t|\theta_i, h^{t-1}\right) = \sum_{\theta'_{-i}} \mu_i(\theta'_{-i}|\theta_i, h^{t-1})\sigma_{-i}\left(a_{-i}^{t+1}|\theta'_{-i}, h^{t-1}\right).$$

In QCSE, there is a parameter $\chi \in [0,1]$. For any $\chi$, $\chi$-cursed player $i$ believes the other players are playing the behavioral strategy:

$$\sigma_{-i}^{\chi}(a_{-i}^t|\theta, h^{t-1}) = \chi\bar{\sigma}_{-i}(a_{-i}^t|\theta_i, h^{t-1}) + (1-\chi)\sigma_{-i}(a_{-i}^t|\theta_{-i}, h^{t-1}),$$

which is a linear combination between the average behavioral strategy (with $\chi$ weight) and the true behavioral strategy (with $1-\chi$ weight). When $\chi = 0$, players have correct perceptions about others' behavioral strategies. On the other extreme, when $\chi = 1$, players fail to understand the correlation between others' actions and types. As the game progresses, players update their beliefs via Bayes' rule, believing that other players are using $\sigma_{-i}^{\chi}$ instead of the true behavioral strategy $\sigma_{-i}$. As shown by Fong et al. (2025), at any history $h^t = (h^{t-1}, a^t)$, player $i$'s $\chi$-cursed belief is

$$\mu_i^{\chi}(\theta_{-i}|\theta_i, h^t) = \chi\mu_i^{\chi}(\theta_{-i}|\theta_i, h^{t-1}) + (1-\chi)\left[\frac{\mu_i^{\chi}(\theta_{-i}|\theta_i, h^{t-1})\sigma_{-i}(a_{-i}^t|\theta_{-i}, h^{t-1})}{\sum_{\theta'_{-i}}\mu_i^{\chi}(\theta'_{-i}|\theta_i, h^{t-1})\sigma_{-i}(a_{-i}^t|\theta'_{-i}, h^{t-1})}\right],$$

which is a linear combination between the belief from the previous period (with $\chi$ weight) and the Bayesian belief (with $1-\chi$ weight).

Moreover, in QCSE, players make quantal responses rather than best responses. In particular, players make make logit quantal responses, and the precision is determined by a parameter $\lambda \in [0, \infty)$. Consider any information set $\mathcal{I}_i$. For any $a_i \in A_i(\mathcal{I}_i)$, let $\bar{u}_{a_i}$ denote the continuation value of $a_i$ in QCSE. The choice probability of $a_i$ is given by a multinomial logit distribution:

$$\sigma_i(a_i|\mathcal{I}_i) = \frac{e^{\lambda\bar{u}_{a_i}}}{\sum_{a'\in A_i(\mathcal{I}_i)} e^{\lambda\bar{u}_{a'}}}.$$

When $\lambda = 0$, players become insensitive to the payoffs, behaving like level 0 players. As $\lambda$ increases, players' behavior becomes more sensitive to the payoffs. In the limit as $\lambda \to \infty$, players become fully rational and make best responses. In summary, QCSE relaxes the requirements of best response and Bayesian inferences with two parameters, $\lambda \in [0, \infty)$ and $\chi \in [0, 1]$.

**Remark 3.** *When $\chi = 0$, QCSE reduces to AQRE, and as $\lambda \to \infty$, it reduces to CSE.*

To enable a fair comparison between DCH and QCSE, I estimate a Quantal DCH model

(QDCH) where the prior distribution of levels follows Poisson($\tau$), and all levels ($k \geq 1$) of players make logit quantal responses instead of best responses. In essence, Quantal DCH relaxes the requirements of best response and mutual consistency with two parameters, $\lambda \in [0, \infty)$ and $\tau \in [0, \infty)$. A description of QDCH can be found in Appendix B.2.

**Remark 4.** *When $\lambda \to \infty$, QDCH reduces to DCH.*

In addition to QDCH and QCSE, I also estimate DCH and AQRE, which are nested within QDCH and QCSE, respectively.[31] These models are estimated using maximum likelihood estimation, and the construction of the likelihood functions can be found in Appendix B.2. Table 3 presents the estimation results for both the sequential and simultaneous treatments.[32] The comparison between the models is summarized in Figure 8.

Table 3: Estimation Results for the Sequential and the Simultaneous Treatment

| | | Sequential Treatment | | | | Simultaneous Treatment | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | QDCH | DCH | QCSE | AQRE | QDCH | DCH | QCSE | AQRE |
| Parameters | $\lambda$ | 12.371 | | 5.672 | 5.484 | 1774.5 | | 8.740 | 3.839 |
| | S.E. | (2.062) | | (0.821) | (0.426) | — | | — | (0.452) |
| | $\tau$ | 1.309 | 0.277 | | | 0.388 | 0.389 | | |
| | S.E. | (0.220) | (0.043) | | | — | (0.015) | | |
| | $\chi$ | | | 0.101 | | | | 1.000 | |
| | S.E. | | | (0.364) | | | | — | |
| Fitness | LL | -634.72 | -671.78 | -648.36 | -648.40 | -1100.76 | -1100.76 | -1167.68 | -1211.08 |
| | AIC | 1273.43 | 1345.57 | 1300.73 | 1298.80 | 2205.51 | 2203.52 | 2339.37 | 2424.16 |
| | BIC | 1283.29 | 1350.50 | 1310.59 | 1303.73 | 2214.61 | 2208.07 | 2348.46 | 2428.70 |

Comparing these four models, we first observe that in both the sequential and simultaneous treatments, QDCH fits the data significantly better than QCSE (Sequential: Vuong Test p-value = 0.0056; Simultaneous: Vuong Test p-value $< 0.0001$). Without relaxing the best response requirement, DCH's fitness is significantly better than QCSE in the simultaneous treatment (Vuong Test p-value $< 0.0001$). However, in the sequential treatment, QCSE fits the data significantly better than DCH (Vuong Test p-value = 0.0245).

QDCH outperforms other models in both treatments, indicating that the observed violation of strategic equivalence is primarily due to the relaxation of mutual consistency. However, there is evidence of the violation of other behavioral biases. In the sequential treatment, a significant quantal response effect is observed (QDCH vs. DCH: Likelihood Ratio Test p-value $< 0.0001$), but not in the simultaneous treatment (Likelihood Ratio Test p-value = 0.9340). Furthermore, in the simultaneous treatment, a significant cursed effect

[31] CSE cannot be estimated independently as it lacks an error structure in the model.

[32] In the simultaneous treatment, due to the flatness of the log-likelihood functions for both QDCH and QCSE at the MLE estimates, the square roots of the inverse Hessian matrices are not well-defined.

is detected ($\hat{\chi} = 1.000$, p-value $< 0.0001$), but not in the sequential treatment ($\hat{\chi} = 0.101$, p-value $= 0.7846$). This suggests that, players struggle to accurately understand how other players' actions depend on their private information and update their beliefs accordingly in the simultaneous treatment, but not in the sequential treatment.

Lastly, it's worth noting that DCH estimates a significantly lower $\hat{\tau}$ in the sequential treatment compared to the simultaneous treatment (Sequential: $\hat{\tau} = 0.277$; Simultaneous: $\hat{\tau} = 0.389$). In contrast, when introducing quantal responses into DCH, we observe a significantly higher $\hat{\tau}$ in the sequential treatment compared to the simultaneous treatment (Sequential: $\hat{\tau} = 1.309$; Simultaneous: $\hat{\tau} = 0.389$). This suggests that in the simultaneous treatment, all of the randomness can be attributed to level 0 behavior, whereas in the sequential treatment, some randomness is attributable to the mistakes of higher-level players.

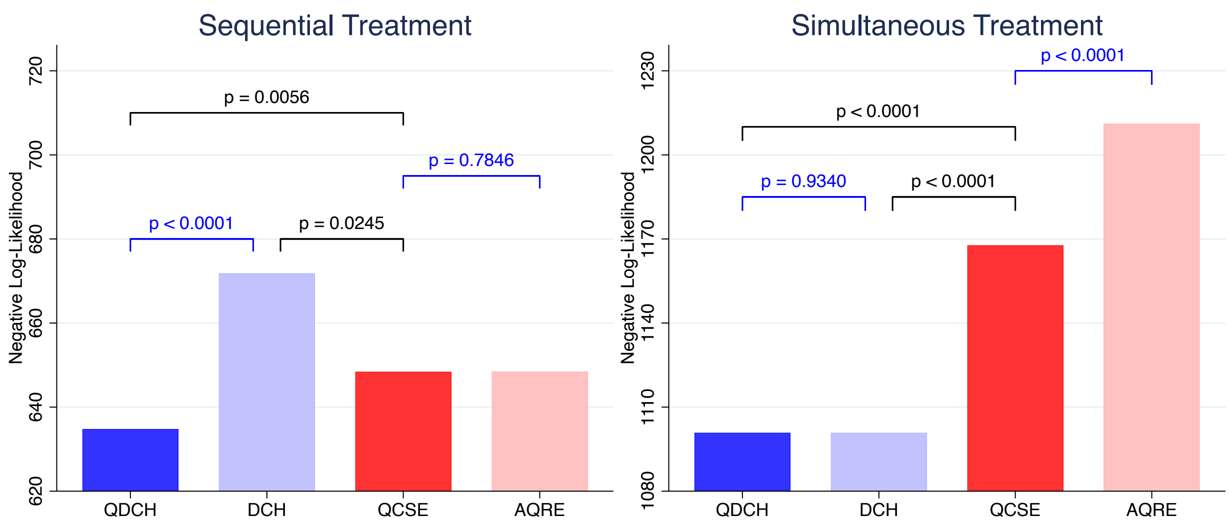



Figure 8: Negative log-likelihood of each model. The likelihood ratio test is performed when comparing two nested models, while the Vuong test is performed when comparing QDCH and DCH with QCSE.

**Result 4.** *(1) In both the sequential and simultaneous treatments, QDCH outperforms QCSE in explaining the data. Additionally, in the sequential treatment, the fitness of DCH is not significantly different from QCSE and AQRE. In the simultaneous treatment, DCH significantly outperforms QCSE and AQRE. (2) In both treatments, there is evidence of the failure of Bayesian inferences. Additionally, in the sequential treatment, evidence of quantal responses is present, while it is not observed in the simultaneous treatment.*

## 5.5 The Analysis of Reaction Times

Besides the choice data, it is also interesting to see how long it takes individuals to make decisions. There is evidence suggesting that people tend to take an action faster if they adopt some simple decision-making heuristics or have strong preferences over the action (see, for example, Rubinstein (2007); Chabris et al. (2009); Konovalov and Krajbich (2019); Lin et al. (2020) and Gill and Prowse (2023)).

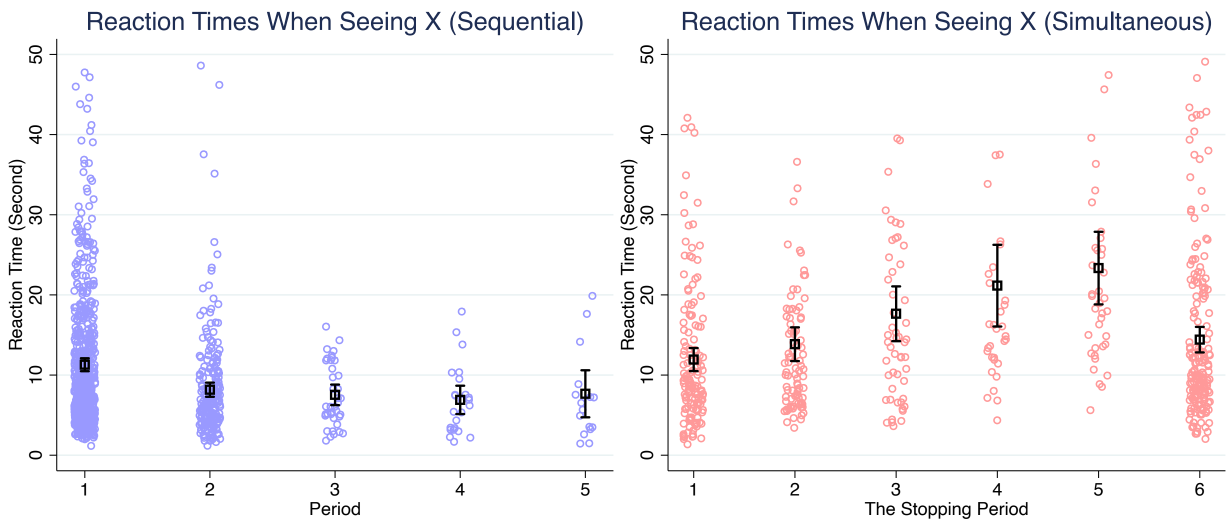



Figure 9: (Left) The reaction time in the sequential treatment. The scatterplot of reaction time conditional on the current period is shown by the blue dots. The mean and the 95% CIs are overlaid. (Right) The reaction time in the simultaneous treatment when seeing a dirty face. The scatterplot of reaction time conditional on the choice of the stopping periods is shown by the red dots. The mean and the 95% CIs are overlaid.

Focusing on the case where players observe a dirty face, Figure 9 presents two panels. The left panel illustrates the distribution of reaction times at each period of the sequential games. The right panel displays the distribution of reaction times for each stopping strategy in the simultaneous games. In the sequential treatment, we can observe that the reaction times at each period are significantly different (Kruskal-Wallis ranksum test: $\chi^2(4) = 32.519$ and p-value $= 0.0001$). Moreover, the reaction time decreases as the game progresses to later periods, dropping from 11.29 seconds in period 1 to 7.66 seconds in period 5. Combined with the low frequencies of $C$ in later periods (around 21.43%), we can conclude that players quickly decide to wait in later periods.

In the simultaneous treatment, we once again observe that the reaction times for each stopping strategy are significantly different (Kruskal-Wallis ranksum test: $\chi^2(5) = 54.291$ and p-value $= 0.0001$). The right panel of Figure 9 reveals a monotonic pattern: players take longer to decide to claim in later periods, with average reaction time of 11.93 seconds for period 1 and 23.34 seconds for period 5. However, it only takes players approximately 14.42 seconds to decide to always wait.

The empirical patterns from both treatments provide suggestive evidence that the heuristic of choosing to "always wait" differs from the heuristic of claiming at a specific period. This finding aligns with the rationale of DCH—level 1 players will always wait upon seeing a dirty face, regardless of the payoff configurations. Conversely, higher-level strategic players will make inferences to determine their stopping strategies. Lastly, the observed monotonic increase in reaction times across stopping strategies in the simultaneous treatment aligns with the idea that choosing to claim at later periods requires more steps of reasoning.

**Result 5.** *(1) In the sequential treatment, when players see a dirty face, their reaction time is shorter in later periods. (2) In the simultaneous treatment, the reaction time of choosing to claim at some period when players see a dirty face is monotonically increasing in the stopping periods. However, players take much less time to decide to always wait.*

# 6 Conclusion

This paper theoretically and experimentally studies the DCH solution, an alternative model that relaxes the mutual consistency requirement of the standard equilibrium theory, in multi-stage games of incomplete information. Instead of mutual consistency, DCH posits that players are heterogeneous with respect to their levels of sophistication and incorrectly believe others are strictly less sophisticated than they are. As the dynamic game progresses, strategic players will update their beliefs about others' types and levels.

In this paper, I characterize some general properties of the DCH belief system in multi-stage games of incomplete information. Proposition 1 guarantees that the DCH belief system is a product measure across players when every player's payoff-relevant type is independently determined. On the other hand, when the prior distribution of types is correlated across players, Proposition 2 demonstrates the existence of a unique corresponding game, where the types are independently drawn, resulting in the DCH solution being invariant in both games. While solving the DCH solution does not require a fixed point argument, it could be computationally challenging in principle, especially when there are more players or information sets involved. To this end, Proposition 1 and 2 simplify the computation, preserving the tractability of DCH. In addition, since strategic players always consider the possibility of others being non-strategic, common knowledge of rationality is never achieved in DCH.

Furthermore, another feature of DCH is the violation of invariance under strategic equivalence, which arises because level 0 players' behavioral strategies are not always outcome-equivalent in different strategically equivalent games, leading to different behavior of higher-level players. To demonstrate the violation of invariance and contrast DCH with the standard equilibrium theory, I characterize the DCH solutions of the sequential and simultaneous two-person dirty-faces games. Despite the two versions of the game sharing the same reduced normal form, the DCH solutions of the two versions differ in a specific way, as characterized by Proposition 5. In summary, DCH predicts that higher-level (level $k \geq 2$) players tend to claim earlier in the sequential version when they are sufficiently impatient, and vice versa in the simultaneous version when they are patient enough.

To test the predictions of DCH, I design and run a laboratory experiment on two-person dirty-faces games where I manipulate both the timing (sequential vs. simultaneous) and payoff structures. The experimental design is guided by DCH, wherein I first calibrate the model using an existing dirty-faces game experimental dataset and choose the payoff parameters to maximize the diagnositicity. Considering that the prior distributions of levels might significantly vary among different subject pools, this experimental design to some extent serves as a stress test for assessing the external validity of DCH.

Overall, a significant treatment effect is detected: players tend to claim earlier in the sequential treatment than in the simultaneous treatment. Some interesting patterns emerge from the data. First, players' behavior significantly varies with payoff structures in both treatments, aligning with the qualitative predictions of DCH. Second, players take longer to choose higher-level stopping strategies. Third, when comparing the fitness of different behavioral models, we find that QDCH outperforms other behavioral models in both treatments. Moreover, there is some evidence of the failure of best responses and Bayesian inferences.

Lastly, when comparing the observed treatment effects with the predictions of the calibrated DCH, we find in one of the diagnostic games where $(\delta, \bar{\alpha}) = (0.6, 0.45)$, the empirical frequency of choosing to claim at period 2 with the observation of a dirty face is 30.77% higher in the sequential treatment than in the simultaneous treatment, which is highly consistent with the calibrated DCH (approximately 31.15%). However, in all control games and the other diagnostic game, the treatment effect is significantly higher than the predictions of the calibrated DCH. Along with the estimation results, we can conclude that while the observed violation of invariance in the data is primarily attributed to the relaxation of mutual consistency, it is a joint consequence of the relaxation of all equilibrium requirements.

The key contribution of this paper is establishing the theoretical and empirical foundations of the DCH solution. However, there are considerable extensions and applications that might be fruitful for future research. The first extension worth pursuing is to endogenize the levels of sophistication, possibly using the cost-benefit analysis proposed by Alaoui and

Penta (2016). This extension could be challenging in dynamic games because each player's level might vary in different information sets of the same game. Additionally, players not only form beliefs about others' current levels but also about their cognitive bounds, which might make the model less tractable.

Second, while the assumption of uniform randomization of level 0 players has some distinct advantages, exploring the actual behavior of level 0 players is another direction worth investigating. Inspired by Li and Camerer (2022), an alternative assumption for level 0 players is that they will randomize across *visually salient* actions at each information set. In particular, due to the rapid development of machine learning algorithms, how visually salient an action is can be quantified even before any behavioral data is collected.

Finally, this last section lists several potential applications of dynamic games of incomplete information where the mutual consistency requirement is easily violated and the DCH solution might provide some new insights.

1. *Social learning*: In social learning games with repeated actions, players make inferences about the true state based on their private signals and publicly observed actions (see Bala and Goyal (1998) and Harel et al. 2021). The DCH solution posits that players do not commonly believe others are able to make correct inferences. Specifically, level 0 players' actions do not convey any information about the true states, while level 1 players will always obey their private signals. For higher-level players, they will constantly update their beliefs about the true state and other players' levels of sophistication. An open question is whether higher-level players will eventually learn the true state.

2. *Sequential bargaining*: The equilibrium of a sequential bargaining game was first characterized by Rubinstein (1982). To reach the perfect equilibrium, players are required to choose the optimal proposal among a continuum of choices at every subgame, and believe the other player to optimally respond to each proposal. Later, McKelvey and Palfrey (1993, 1995) considered a two-person multi-stage bargaining game where each players has a private payoff-relevant type and makes a binary decide (whether to give in or hold out) in every period. The game continues until at least one of the players gives in. In this game, it is strictly dominant for the strong type of players to hold out forever, but not for the weak type—the weak type players need to trade-off between the reward of giving in earlier and the reputational benefit from mimicking the high type. This reasoning is behaviorally challenging. In contrast, DCH is not a solution of a fixed point problem but solved iteratively from lower to higher levels. Therefore, the DCH solution is expected to be sharply different from the standard equilibrium in the sequential bargaining game.

3. *Signaling:* In a multi-stage signaling game, an informed player will have a persistent type and interact with an uninformed player repeatedly. Kaya (2009) analyzed such an environment, finding that the set of equilibrium signal sequences includes a large class of possibly complex signal sequences. In contrast, in the DCH solution, the uninformed player will learn about the informed player's true type and level when observing a new signal, and the informed player will also learn about the uninformed player's level at each stage. Given that the DCH solution is unique, it would be interesting to characterize the signal sequence of each level of informed players and test whether this is consistent with the behavioral data.

4. *Sequential voting:* There is a large class of voting rules that includes multiple rounds, such as sequential voting over agendas (Baron and Ferejohn, 1989) or elections based on repeated ballots and elimination of one candidate in each round (Bag et al., 2009). To reach Condorcet consistent outcomes, players are required to behave strategically. However, in cases where voters are not strategic or believe others might not be strategic, the DCH solution becomes an ideal solution concept. In the DCH solution, voters will update their beliefs about others' preferences and levels of sophistication simultaneously, and vote according to their posterior beliefs in each round. Since the common knowledge of rationality is violated in DCH, it is natural to conjecture that higher-level players will vote more sincerely in DCH than in the equilibrium.

# References


Alaoui, L. and A. Penta (2016): "Endogenous depth of reasoning," *The Review of Economic Studies*, 83, 1297–1333.

An, Y. (2017): "Identification of First-Price Auctions with Non-Equilibrium Beliefs: A Measurement Error Approach," *Journal of Econometrics*, 200, 326–343.

Bag, P. K., H. Sabourian, and E. Winter (2009): "Multi-stage voting, sequential elimination and Condorcet consistency," *Journal of Economic Theory*, 144, 1278–1299.

Bala, V. and S. Goyal (1998): "Learning from neighbours," *The review of economic studies*, 65, 595–621.

Baron, D. P. and J. A. Ferejohn (1989): "Bargaining in legislatures," *American political science review*, 83, 1181–1206.

Barwise, J. (1981): "Scenes and Other Situations," *Journal of Philosophy*, 78, 369–397.

Battigalli, P. (2024): "Reduced Strategies and Cognitive Hierarchies in the Extensive and Normal Form," *Available at SSRN 5064371*.

Bayer, R. C. and M. Chan (2007): "The Dirty Faces Game Revisited," *Working Paper*.

Bayer, R. C. and L. Renou (2016a): "Logical Abilities and Behavior in Strategic-Form Games," *Journal of Economic Psychology*, 56, 39–59.

——— (2016b): "Logical Omniscience at the Laboratory," *Journal of Behavioral and Experimental Economics*, 64, 41–49.

Bersson, E., P. Hürtgen, and M. Paustian (2024): "Expectations Formation, Sticky Prices, and the ZLB," *Journal of Money, Credit and Banking*, 56, 365–393.

Binmore, K. and A. Brandeburger (1988): "Common Knowledge and Game Theory," *MichU DeptE CenREST W89-06*.

Brandts, J. and G. Charness (2011): "The Strategy Versus the Direct-Response Method: A First Survey of Experimental Comparisons," *Experimental Economics*, 14, 375–398.

Brown, A. L., C. F. Camerer, and D. Lovallo (2012): "To Review or Not to Review? Limited Strategic Thinking at the Movie Box Office," *American Economic Journal: Microeconomics*, 4, 1–26.

Cai, H. and J. T.-y. Wang (2006): "Overcommunication in Strategic Information Transmission Games," *Games and Economic Behavior*, 56, 7–36.

Camerer, C. and D. Lovallo (1999): "Overconfidence and Excess Entry: An Experimental Approach," *American Economic Review*, 89, 306–318.

Camerer, C. F. (2003): *Behavioral Game Theory: Experiments in Strategic Interaction*, Princeton University Press.

Camerer, C. F., T.-H. Ho, and J.-K. Chong (2004): "A Cognitive Hierarchy Model of Games," *Quarterly Journal of Economics*, 119, 861–898.

Chabris, C. F., C. L. Morris, D. Taubinsky, D. Laibson, and J. P. Schuldt (2009): "The allocation of time in decision-making," *Journal of the European Economic Association*, 7, 628–637.

Chen, D. L., M. Schonger, and C. Wickens (2016): "oTree—An open-source platform for laboratory, online, and field experiments," *Journal of Behavioral and Experimental Finance*, 9, 88–97.

Chen, W. J., M.-J. Fong, and P.-H. Lin (2024): "Measuring Higher-Order Rationality with Belief Control," *arXiv preprint arXiv:2309.07427v2*.

Choo, L. and X. Zhou (2022): "Can Market Selection Reduce Anomalous Behaviour in Games?" *European Economic Review*, 141, 103958.

Cohen, S. and S. Li (2022): "Sequential Cursed Equilibrium," *arXiv preprint arXiv:2212.06025*.

Coricelli, G. and R. Nagel (2009): "Neural Correlates of Depth of Strategic Reasoning in Medial Prefrontal Cortex," *Proceedings of the National Academy of Sciences*, 106, 9163–9168.

Costa-Gomes, M., V. P. Crawford, and B. Broseta (2001): "Cognition and Behavior in Normal-Form Games: An Experimental Study," *Econometrica*, 69, 1193–1235.

Costa-Gomes, M. A. and V. P. Crawford (2006): "Cognition and Behavior in Two-Person Guessing Games: An Experimental Study," *American Economic Review*, 96, 1737–1768.

Crawford, V. P. and N. Iriberri (2007a): "Fatal Attraction: Salience, Naivete, and Sophistication in Experimental "Hide-and-Seek" Games," *American Economic Review*, 97, 1731–1750.

——— (2007b): "Level-$k$ auctions: Can a Nonequilibrium Model of Strategic Thinking Explain the Winner's Curse and Overbidding in Private-Value Auctions?" *Econometrica*, 75, 1721–1770.

Devetag, G. and M. Warglien (2003): "Games and Phone Numbers: Do Short-Term Memory Bounds Affect Strategic Behavior?" *Journal of Economic Psychology*, 24, 189–202.

Elmes, S. and P. J. Reny (1994): "On the strategic equivalence of extensive form games," *Journal of Economic Theory*, 62, 1–23.

Enke, B., T. Graeber, and R. Oprea (2023): "Confidence, Self-Selection, and Bias in the Aggregate," *American Economic Review*, 113, 1933–1966.

Fong, M.-J., P.-H. Lin, and T. R. Palfrey (2023): "A Note on Cursed Sequential Equilibrium and Sequential Cursed Equilibrium," *arXiv preprint arXiv:2304.05515*.

——— (2025): "Cursed Sequential Equilibrium," *American Economic Review*, 115, 2616–2658.

Fudenberg, D. and J. Tirole (1991): "Perfect Bayesian Equilibrium and Sequential Equilibrium," *Journal of Economic Theory*, 53, 236–260.

——— (1993): *Game Theory*, MIT press.

Gamow, G. and M. Stern (1958): "Forty Unfaithful Wives," *Puzzle Math*, 20–23.

García-Pola, B., N. Iriberri, and J. Kovářík (2020): "Hot versus Cold Behavior in Centipede Games," *Journal of the Economic Science Association*, 6, 226–238.

Geanakoplos, J. (1994): "Common Knowledge," *Handbook of Game Theory with Economic Applications*, 2, 1437–1496.

Gill, D. and V. Prowse (2023): "Strategic complexity and the value of thinking," *The Economic Journal*, 133, 761–786.

Gillen, B. (2009): "Identification and Estimation of Level-$k$ Auctions," *Available at SSRN 1337843*.

Goldfarb, A. and M. Xiao (2011): "Who Thinks about the Competition? Managerial Ability and Strategic Entry in US Local Telephone Markets," *American Economic Review*, 101, 3130–3161.

Gong, B. and Y. Liang (2024): "A Dynamic Matching Mechanism for College Admissions: Theory and Experiment," *Management Science*.

Grehl, S. and A. Tutić (2015): "Experimental Evidence on Iterated Reasoning in Games," *PloS One*, 10, e0136524.

Halpern, J. Y. and Y. Moses (1990): "Knowledge and Common Knowledge in a Distributed Environment," *Journal of the ACM (JACM)*, 37, 549–587.

Hardin, C. S. and A. D. Taylor (2008): "An Introduction to Infinite Hat Problems," *The Mathematical Intelligencer*, 30, 20–25.

Harel, M., E. Mossel, P. Strack, and O. Tamuz (2021): "Rational groupthink," *The Quarterly Journal of Economics*, 136, 621–668.

Hortaçsu, A., F. Luco, S. L. Puller, and D. Zhu (2019): "Does Strategic Ability Affect Efficiency? Evidence from Electricity Markets," *American Economic Review*, 109, 4302–4342.

Kaya, A. (2009): "Repeated signaling games," *Games and Economic Behavior*, 66, 841–854.

Kohlberg, E. and J.-F. Mertens (1986): "On the Strategic Stability of Equilibria," *Econometrica*, 1003–1037.

Konovalov, A. and I. Krajbich (2019): "Revealed strength of preference: Inference from response times," *Judgment and Decision making*, 14, 381–394.

Li, S. (2017): "Obviously Strategy-Proof Mechanisms," *American Economic Review*, 107, 3257–3287.

Li, X. and C. F. Camerer (2022): "Predictable effects of visual salience in experimental decisions and games," *The Quarterly Journal of Economics*, 137, 1849–1900.

Lin, P.-H. (2023): "Cognitive Hierarchies in Multi-Stage Games of Incomplete Information: Theory and Experiment," *arXiv preprint arXiv:2208.11190v3*.

Lin, P.-H., A. L. Brown, T. Imai, J. T.-y. Wang, S. W. Wang, and C. F. Camerer (2020): "Evidence of general economic principles of bargaining and trade from 2,000 classroom experiments," *Nature Human Behaviour*, 4, 917–927.

Lin, P.-H. and T. R. Palfrey (2024): "Cognitive Hierarchies for Games in Extensive Form," *Journal of Economic Theory*, 105871.

Littlewood, J. E. (1953): *A Mathematician's Miscellany*, London, England: Meuthen & Co. Ltd.

Liu, Z. (2008): "The Dirty Face Problem with Unawareness," *The BE Journal of Theoretical Economics*, 8, 1935–1704.

McKelvey, R. D. and T. R. Palfrey (1993): "Engodeneity of Alternating Offers in a Bargaining Game," *Caltech Social Science Working Paper*.

——— (1995): "The holdout game: An experimental study of an infinitely repeated game with," in *Social choice, welfare, and ethics: Proceedings of the Eighth International Symposium in Economic Theory and Econometrics*, Cambridge University Press, vol. 8, 321.

——— (1998): "Quantal Response Equilibria for Extensive Form Games," *Experimental Economics*, 1, 9–41.

Moore, D. A. and P. J. Healy (2008): "The Trouble with Overconfidence," *Psychological Review*, 115, 502.

Moses, Y., D. Dolev, and J. Y. Halpern (1986): "Cheating Husbands and Other Stories: A Case Study of Knowledge, Action, and Communication," *Distributed Computing*, 1, 167–176.

Myerson, R. B. (1985): "Bayesian Equilibrium and Incentive-Compatibility: An Introduction," *Social Goals and Social Organization: Essays in Memory of Elisha Pazner*, 229–260.

Nagel, R. (1995): "Unraveling in Guessing Games: An Experimental Study," *American Economic Review*, 85, 1313–1326.

Östling, R., J. T.-y. Wang, E. Y. Chou, and C. F. Camerer (2011): "Testing Game Theory in the Field: Swedish LUPI Lottery Games," *American Economic Journal: Microeconomics*, 3, 1–33.

Richie, R., J. Seitz-Brown, and L. Kaufman (2023): "The Case for Instant Runoff Voting," *Constitutional Political Economy*, 34, 367–377.

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.

Rubinstein, A. (1982): "Perfect equilibrium in a bargaining model," *Econometrica: Journal of the Econometric Society*, 97–109.

——— (2007): "Instinctive and cognitive reasoning: A study of response times," *The Economic Journal*, 117, 1243–1259.

Schipper, B. C. and H. Zhou (2024): "Level-$k$ Thinking in the Extensive Form," *Economic Theory*, 1–41.

Selten, R. (1967): "Die Strategiemethode zur Erforschung des eingeschränkt rationalen Verhaltens im Rahmen eines Oligopolexperiments," in *Beiträge zur experimentellen Wirtschaftsforschung*, Tübingen: Mohr, 136–168.

Stahl, D. O. and P. W. Wilson (1994): "Experimental Evidence on Players' Models of Other Players," *Journal of Economic Behavior & Organization*, 25, 309–327.

——— (1995): "On Players' Models of Other Players: Theory and Experimental Evidence," *Games and Economic Behavior*, 10, 218–254.

Thompson, F., R. Corporation, and S. U. d'America. Air force. Project Rand (1952): *Equivalence of Games in Extensive Form*, Research memorandum / U.S. Air force, Rand Corporation.

Wald, A. (1947): *Sequential analysis*, John Wiley and Sons, New York.

Wang, J. T.-y., M. Spezio, and C. F. Camerer (2010): "Pinocchio's Pupil: Using Eyetracking and Pupil Dilation to Understand Truth Telling and Deception in Sender-Receiver Games," *American Economic Review*, 100, 984–1007.

Weber, R. A. (2001): "Behavior and Learning in the "Dirty Faces" Game," *Experimental Economics*, 4, 229–242.

# Appendix

## Proof of Proposition 1

To prove this proposition, I first characterize the posterior beliefs in Lemma 1 then prove that the beliefs are independent across players if the types are independently drawn.

**Lemma 1.** *Consider any multi-stage game with observed actions $\Gamma$, any $i \in N$, $\theta_i \in \Theta_i$, $h \in \mathcal{H} \backslash \mathcal{H}^T$, and every level $k \in \mathbb{N}$. For every information set $I_i = (\theta_i, h)$, level-k player i's belief at $I_i$ can be characterized as follows.*

1. *Level-k player i's prior belief about other players' types and levels are independent. That is, $\mu^k(\theta_{-i}, \tau_{-i}|\theta_i, h_\emptyset) = F(\theta_{-i}|\theta_i) \prod_{j \neq i} \hat{P}^k_{ij}(\tau_j)$.*
2. *For any $1 \leq t < T$, and $h^t \in \mathcal{H}^t$, level-k player i's belief at information set $(\theta_i, h^t)$ where $h^t = (a^1, \ldots, a^t)$ is*

$$\mu^k(\theta_{-i}, \tau_{-i}|\theta_i, h^t) = \frac{F(\theta_{-i}|\theta_i) \prod_{j \neq i} \left\{ \hat{P}^k_{ij}(\tau_j) \prod_{l=1}^t \sigma_j^{\tau_j}(a^l_j|\theta_j, h^{l-1}) \right\}}{\sum_{\theta'_{-i}} \sum_{\{\tau'_{-i}: \tau'_j < k \ \forall j \neq i\}} F(\theta'_{-i}|\theta_i) \prod_{j \neq i} \left\{ \hat{P}^k_{ij}(\tau'_j) \prod_{l=1}^t \sigma_j^{\tau'_j}(a^l_j|\theta'_j, h^{l-1}) \right\}}.$$

**Proof of Lemma 1:**

***1.*** At the beginning of the game, the only information available to player $i$ is his own type $\theta_i$ and his level of sophistication $\tau_i = k$. Therefore, the prior belief is the probability of the opponents' types and levels conditional on $\theta_i$ and $\tau_i$, which is

$$\mu^k(\theta_{-i}, \tau_{-i}|\theta_i, h_\emptyset) = \Pr(\theta_{-i}, \tau_{-i}|\theta_i, \tau_i = k) = \Pr(\theta_{-i}|\theta_i)\Pr(\tau_{-i}|\tau_i = k) = F(\theta_{-i}|\theta_i)\prod_{j\neq i}\hat{P}^k_{ij}(\tau_j).$$

The second equality holds because the types and levels are independently drawn.

***2.*** This can be shown by induction on $t$. Consider any available history at period 2, $h^1 \in \mathcal{H}^1$. Level-$k$ player $i$'s belief at information set $(\theta_i, h^1)$ is

$$\mu^k(\theta_{-i}, \tau_{-i}|\theta_i, h^1) = \frac{\mu^k(\theta_{-i}, \tau_{-i}|\theta_i, h_\emptyset)\prod_{j\neq i}\sigma_j^{\tau_j}(a_j^1|\theta_j, h_\emptyset)}{\sum_{\theta'_{-i}}\sum_{\{\tau'_{-i}:\tau'_j<k\ \forall j\neq i\}}\mu^k(\theta'_{-i}, \tau'_{-i}|\theta_i, h_\emptyset)\prod_{j\neq i}\sigma_j^{\tau'_j}(a_j^1|\theta'_j, h_\emptyset)}. \quad \text{(A.1)}$$

By step 1, we know $\mu^k(\theta_{-i}, \tau_{-i}|\theta_i, h_\emptyset) = F(\theta_{-i}|\theta_i)\prod_{j\neq i}\hat{P}^k_{ij}(\tau_j)$. Plugging in Equation (A.1), we can obtain that

$$\begin{aligned}\mu^k(\theta_{-i}, \tau_{-i}|\theta_i, h^1) &= \frac{\mu^k(\theta_{-i}, \tau_{-i}|\theta_i, h_\emptyset)\prod_{j\neq i}\sigma_j^{\tau_j}(a_j^1|\theta_j, h_\emptyset)}{\sum_{\theta'_{-i}}\sum_{\{\tau'_{-i}:\tau'_j<k\ \forall j\neq i\}}\mu^k(\theta'_{-i}, \tau'_{-i}|\theta_i, h_\emptyset)\prod_{j\neq i}\sigma_j^{\tau'_j}(a_j^1|\theta'_j, h_\emptyset)} \\ &= \frac{F(\theta_{-i}|\theta_i)\prod_{j\neq i}\left\{\hat{P}^k_{ij}(\tau_j)\sigma_j^{\tau_j}(a_j^1|\theta_j, h_\emptyset)\right\}}{\sum_{\theta'_{-i}}\sum_{\{\tau'_{-i}:\tau'_j<k\ \forall j\neq i\}}F(\theta'_{-i}|\theta_i)\prod_{j\neq i}\left\{\hat{P}^k_{ij}(\tau'_j)\sigma_j^{\tau'_j}(a_j^1|\theta'_j, h_\emptyset)\right\}}.\end{aligned}$$

Next, suppose there is $t'$ such that the statement holds for every period $t = 2, \ldots, t'$. Consider period $t'+1$ and any history available at period $t'+1$, $h^{t'} \in \mathcal{H}^{t'}$ where $h^{t'} = (a^1, \ldots, a^{t'})$. Then level-$k$ player $i$'s belief at information set $(\theta_i, h^{t'})$ is

$$\begin{aligned}\mu^k(\theta_{-i}, \tau_{-i}|\theta_i, h^{t'}) &= \frac{\mu^k(\theta_{-i}, \tau_{-i}|\theta_i, h^{t'-1})\prod_{j\neq i}\sigma_j^{\tau_j}(a_j^{t'}|\theta_j, h^{t'-1})}{\sum_{\theta'_{-i}}\sum_{\{\tau'_{-i}:\tau'_j<k\ \forall j\neq i\}}\mu^k(\theta'_{-i}, \tau'_{-i}|\theta_i, h^{t'-1})\prod_{j\neq i}\sigma_j^{\tau'_j}(a_j^{t'}|\theta'_j, h^{t'-1})} \\ &= \frac{F(\theta_{-i}|\theta_i)\prod_{j\neq i}\left\{\hat{P}^k_{ij}(\tau_j)\prod_{l=1}^{t'-1}\sigma_j^{\tau_j}(a_j^l|\theta_j, h^{l-1})\right\}\prod_{j\neq i}\sigma_j^{\tau_j}(a_j^{t'}|\theta_j, h^{t'-1})}{\sum_{\theta'_{-i}}\sum_{\{\tau'_{-i}:\tau'_j<k\ \forall j\neq i\}}F(\theta'_{-i}|\theta_i)\prod_{j\neq i}\left\{\hat{P}^k_{ij}(\tau'_j)\prod_{l=1}^{t'-1}\sigma_j^{\tau'_j}(a_j^l|\theta'_j, h^{l-1})\right\}\prod_{j\neq i}\sigma_j^{\tau'_j}(a_j^{t'}|\theta'_j, h^{t'-1})} \\ &= \frac{F(\theta_{-i}|\theta_i)\prod_{j\neq i}\left\{\hat{P}^k_{ij}(\tau_j)\prod_{l=1}^{t'}\sigma_j^{\tau_j}(a_j^l|\theta_j, h^{l-1})\right\}}{\sum_{\theta'_{-i}}\sum_{\{\tau'_{-i}:\tau'_j<k\ \forall j\neq i\}}F(\theta'_{-i}|\theta_i)\prod_{j\neq i}\left\{\hat{P}^k_{ij}(\tau'_j)\prod_{l=1}^{t'}\sigma_j^{\tau'_j}(a_j^l|\theta'_j, h^{l-1})\right\}}.\end{aligned}$$

The second equality holds because of the induction hypothesis, as desired. $\square$

**Proof of Proposition 1:**

We prove this by induction on $t$. Let $\sigma$ be any level-dependent strategy profile and $F$ and $P$ be any distributions of types and levels. First, consider $t = 1$. By Lemma 1, we know $\mu^k(\theta_{-i}, \tau_{-i}|\theta_i, h_\emptyset) = F(\theta_{-i}|\theta_i)\prod_{j\neq i}\hat{P}^k_{ij}(\tau_j)$. As the prior distribution of types is independent across players, we can obtain that

$$\begin{aligned}\mu^k(\theta_{-i}, \tau_{-i}|\theta_i, h_\emptyset) = F(\theta_{-i}|\theta_i)\prod_{j\neq i}\hat{P}^k_{ij}(\tau_j) &= \prod_{j\neq i}F_j(\theta_j)\prod_{j\neq i}\hat{P}^k_{ij}(\tau_j) \\ &= \prod_{j\neq i}\left[F_j(\theta_j)\hat{P}^k_{ij}(\tau_j)\right] = \prod_{j\neq i}\mu^k_j(\theta_j, \tau_j|\theta_i, h_\emptyset).\end{aligned}$$

Therefore, we know the result is true at $t = 1$. Next, suppose there is $t' > 1$ such that the result holds for all $t = 1, \ldots, t'$. We want to show that the result holds at period $t' + 1$. Let $h^{t'} \in \mathcal{H}^{t'}$ be any available history in period $t' + 1$ where $h^{t'} = (h^{t'-1}, a^{t'})$. Therefore, player $i$'s posterior belief at history $h^{t'}$ is

$$\mu^k(\theta_{-i}, \tau_{-i}|\theta_i, h^{t'}) = \frac{\mu^k(\theta_{-i}, \tau_{-i}|\theta_i, h^{t'-1})\prod_{j\neq i}\sigma_j^{\tau_j}(a_j^{t'}|\theta_j, h^{t'-1})}{\sum_{\theta'_{-i}\in\Theta_{-i}}\sum_{\{\tau'_{-i}:\tau'_j<k\ \forall j\neq i\}}\mu^k(\theta'_{-i}, \tau'_{-i}|\theta_i, h^{t'-1})\prod_{j\neq i}\sigma_j^{\tau'_j}(a_j^{t'}|\theta'_j, h^{t'-1})}.$$

By induction hypothesis, we know $\mu^k(\theta_{-i}, \tau_{-i}|\theta_i, h^{t'-1}) = \prod_{j\neq i}\mu^k_j(\theta_j, \tau_j|\theta_i, h^{t'-1})$. Therefore, as we rearrange the posterior belief $\mu^k(\theta_{-i}, \tau_{-i}|\theta_i, h^{t'})$, we can obtain that

$$\begin{aligned}\mu^k(\theta_{-i}, \tau_{-i}|\theta_i, h^{t'}) &= \frac{\mu^k(\theta_{-i}, \tau_{-i}|\theta_i, h^{t'-1})\prod_{j\neq i}\sigma_j^{\tau_j}(a_j^{t'}|\theta_j, h^{t'-1})}{\sum_{\theta'_{-i}\in\Theta_{-i}}\sum_{\{\tau'_{-i}:\tau'_j<k\ \forall j\neq i\}}\mu^k(\theta'_{-i}, \tau'_{-i}|\theta_i, h^{t'-1})\prod_{j\neq i}\sigma_j^{\tau'_j}(a_j^{t'}|\theta'_j, h^{t'-1})} \\ &= \frac{\prod_{j\neq i}\left[\mu^k_j(\theta_j, \tau_j|\theta_i, h^{t'-1})\sigma_j^{\tau_j}(a_j^{t'}|\theta_j, h^{t'-1})\right]}{\sum_{\theta'_{-i}\in\Theta_{-i}}\sum_{\{\tau'_{-i}:\tau'_j<k\ \forall j\neq i\}}\prod_{j\neq i}\left[\mu^k_j(\theta'_j, \tau'_j|\theta_i, h^{t'-1})\sigma_j^{\tau'_j}(a_j^{t'}|\theta'_j, h^{t'-1})\right]} \\ &= \frac{\prod_{j\neq i}\left[\mu^k_j(\theta_j, \tau_j|\theta_i, h^{t'-1})\sigma_j^{\tau_j}(a_j^{t'}|\theta_j, h^{t'-1})\right]}{\sum_{\theta'_{-i}\in\Theta_{-i}}\prod_{j\neq i}\left[\sum_{\tau'_j<k}\mu^k_j(\theta'_j, \tau'_j|\theta_i, h^{t'-1})\sigma_j^{\tau'_j}(a_j^{t'}|\theta'_j, h^{t'-1})\right]} \\ &= \frac{\prod_{j\neq i}\left[\mu^k_j(\theta_j, \tau_j|\theta_i, h^{t'-1})\sigma_j^{\tau_j}(a_j^{t'}|\theta_j, h^{t'-1})\right]}{\prod_{j\neq i}\left[\sum_{\theta'_j\in\Theta_j}\sum_{\tau'_j<k}\mu^k_j(\theta'_j, \tau'_j|\theta_i, h^{t'-1})\sigma_j^{\tau'_j}(a_j^{t'}|\theta'_j, h^{t'-1})\right]}.\end{aligned}$$

As a result, we can conclude that

$$\begin{aligned}\mu^k(\theta_{-i},\tau_{-i}|\theta_i,h^{t'}) &= \prod_{j\neq i}\left[\frac{\mu_j^k(\theta_j,\tau_j|\theta_i,h^{t'-1})\sigma_j^{\tau_j}(a_j^{t'}|\theta_j,h^{t'-1})}{\sum_{\theta'_j\in\Theta_j}\sum_{\tau'_j<k}\mu_j^k(\theta'_j,\tau'_j|\theta_i,h^{t'-1})\sigma_j^{\tau'_j}(a_j^{t'}|\theta'_j,h^{t'-1})}\right]\\ &= \prod_{j\neq i}\mu_j^k(\theta_j,\tau_j|\theta_i,h^{t'}).\end{aligned}$$

This completes the proof of the proposition. ■

## Proof of Proposition 2

By Lemma 1, we know that in the transformed game (with independent types) $\hat{\Gamma}$, level-$k$ player $i$'s belief at $h^t\in\mathcal{H}^t$ is

$$\begin{aligned}\hat{\mu}^k(\theta_{-i},\tau_{-i}|\theta_i,h^t) &= \frac{\hat{F}(\theta_{-i}|\theta_i)\prod_{j\neq i}\left\{\hat{P}_{ij}^k(\tau_j)\prod_{l=1}^t\sigma_j^{\tau_j}(a_j^l|\theta_j,h^{l-1})\right\}}{\sum_{\theta'_{-i}}\sum_{\{\tau'_{-i}:\tau'_j<k\ \forall j\neq i\}}\hat{F}(\theta'_{-i}|\theta_i)\prod_{j\neq i}\left\{\hat{P}_{ij}^k(\tau'_j)\prod_{l=1}^t\sigma_j^{\tau'_j}(a_j^l|\theta'_j,h^{l-1})\right\}}\\ &= \frac{\prod_{j\neq i}\left\{\hat{P}_{ij}^k(\tau_j)\prod_{l=1}^t\sigma_j^{\tau_j}(a_j^l|\theta_j,h^{l-1})\right\}}{\sum_{\theta'_{-i}}\sum_{\{\tau'_{-i}:\tau'_j<k\ \forall j\neq i\}}\prod_{j\neq i}\left\{\hat{P}_{ij}^k(\tau'_j)\prod_{l=1}^t\sigma_j^{\tau'_j}(a_j^l|\theta'_j,h^{l-1})\right\}}.\end{aligned}$$

Therefore, we can obtain that

$$\begin{aligned}\mu^k(\theta_{-i},\tau_{-i}|\theta_i,h^t) &= \frac{F(\theta_{-i}|\theta_i)\prod_{j\neq i}\left\{\hat{P}_{ij}^k(\tau_j)\prod_{l=1}^t\sigma_j^{\tau_j}(a_j^l|\theta_j,h^{l-1})\right\}}{\sum_{\theta'_{-i}}\sum_{\{\tau'_{-i}:\tau'_j<k\ \forall j\neq i\}}F(\theta'_{-i}|\theta_i)\prod_{j\neq i}\left\{\hat{P}_{ij}^k(\tau'_j)\prod_{l=1}^t\sigma_j^{\tau'_j}(a_j^l|\theta'_j,h^{l-1})\right\}}\\ &= \frac{F(\theta_{-i}|\theta_i)\hat{\mu}^k(\theta_{-i},\tau_{-i}|\theta_i,h^t)}{\sum_{\theta'_{-i}}\sum_{\{\tau'_{-i}:\tau'_j<k\ \forall j\neq i\}}F(\theta'_{-i}|\theta_i)\hat{\mu}^k(\theta'_{-i},\tau'_{-i}|\theta_i,h^t)}.\end{aligned}$$

To complete the proof, it suffices to show that for each level-$k$ player $i$ and every $h^t\in\mathcal{H}\backslash\mathcal{H}^T$, maximizing $\mathbb{E}u_i^k$ given belief $\mu_i^k$ and $\sigma_{-i}^{<k}$ is equivalent to maximizing $\mathbb{E}\hat{u}_i^k$ given belief $\hat{\mu}^k$ and $\hat{\sigma}_{-i}^{<k}=\sigma_{-i}^{<k}$. This is true because the expected payoff in the original game (with correlated types) is:

$$\begin{aligned}&\mathbb{E}u_i^k(\sigma|\theta_i,h^t) =\\ &\qquad\sum_{h^T\in\mathcal{H}^T}\sum_{\theta_{-i}\in\Theta_{-i}}\sum_{\{\tau_{-i}:\tau_j<k\ \forall j\neq i\}}\mu^k(\theta_{-i},\tau_{-i}|\theta_i,h^t)P_i^k(h^T|\theta,h^t,\tau_{-i},\sigma_{-i}^{<k},\sigma_i^k)u_i(\theta_i,\theta_{-i},h^T),\end{aligned}$$

which is proportional to

$$
\begin{aligned}
&\mathbb{E}\hat{u}_i^k(\sigma|\theta_i, h^t) = \\
&\sum_{h^T \in \mathcal{H}^T} \sum_{\theta_{-i} \in \Theta_{-i}} \sum_{\{\tau_{-i}: \tau_j < k\ \forall j \neq i\}} F(\theta_{-i}|\theta_i)\hat{\mu}^k(\theta_{-i}, \tau_{-i}|\theta_i, h^t) P_i^k(h^T|\theta, h^t, \tau_{-i}, \sigma_{-i}^{<k}, \sigma_i^k) u_i(\theta_i, \theta_{-i}, h^T) \\
&= \sum_{h^T \in \mathcal{H}^T} \sum_{\theta_{-i} \in \Theta_{-i}} \sum_{\{\tau_{-i}: \tau_j < k\ \forall j \neq i\}} \hat{\mu}^k(\theta_{-i}, \tau_{-i}|\theta_i, h^t) P_i^k(h^T|\theta, h^t, \tau_{-i}, \sigma_{-i}^{<k}, \sigma_i^k) \hat{u}_i(\theta_i, \theta_{-i}, h^T).
\end{aligned}
$$

This completes the proof of the proposition. ■

# A Detailed Analysis of Bayer and Chan (2007) Data (For Online Publication)

## A.1 Data Description

This section revisits the dirty-faces experimental data by Bayer and Chan (2007). The description of the experimental setting can be found in the main text section 4.1, and the instructions and screenshots can be found in Bayer and Chan (2007) Appendix A.

Following previous notations, I use $(t, x_{-i})$ to denote the situation where subject $i$ sees type $x_{-i}$ at period $t$. After excluding the data from the case where there is no public announcement, the raw data at each information set is reported in Table A.1. Each entry in the table states the number of observations and the percentage of the choices that follow the equilibrium predictions. For instance, at information set $(t, x_{-i}) = (2, X)$, there are 170 choices and 62 percent of the choices are $C$, which is the action predicted by the equilibrium.

Table A.1: Experimental Data from Bayer and Chan (2007)

| | Number of Players | | | | |
|---|---|---|---|---|---|
| | 2 | | 3 | | |
| $x_{-i}$ | $O$ | $X$ | $OO$ | $OX$ | $XX$ |
| EQ | $C$ | $WC$ | $C$ | $WC$ | $WWC$ |
| Period | Number of Obs (EQ %) | | | | |
| 1 | 123 (0.94) | 391 (0.79) | 48 (0.92) | 280 (0.61) | 320 (0.76) |
| 2 | 6 (0.50) | 170 (0.62) | 2 (0.50) | 60 (0.58) | 145 (0.79) |
| 3 | — | — | — | 10 (0.20) | 56 (0.36) |

Note: In Treatment 1, there are 21 groups of subjects (42 subjects in total), and in Treatment 2, there are 16 groups of subjects (48 subjects in total). Because each group plays 14 rounds, the data set consists of $(21 + 16) \times 14 = 518$ games.

From Table A.1, we can observe that the behavior aligns with the equilibrium prediction when players do not see any dirty face. In this situation ($x_{-i} = O$ or $OO$), players are aware that their face type is $X$ and choose $C$ in period 1. However, the behavior becomes less consistent with the equilibrium as the reasoning complexity increases. When players see only one dirty face ($x_{-i} = X$ or $OX$), they should realize that their face type is $X$ as the game progresses to period 2. However, the empirical data show that only 62% and 58% of players in Treatment 1 and 2, respectively, are able to do so. Furthermore, when players see two dirty faces ($x_{-i} = XX$), only 30% of the players claim to have a dirty face in period 3.

These observations suggest that the equilibrium fails to explain a significant portion of the data. In the following analysis, I compare the fitness of the DCH model with that of the standard CH model and the agent quantal response equilibrium (AQRE) proposed by

McKelvey and Palfrey (1998). By comparing the DCH and the standard CH models, I can quantify the improvement achieved by incorporating learning from past actions into the CH framework. On the other hand, AQRE is an equilibrium model designed for extensive games, where players make stochastic choices and assume that other players do the same. The comparison between the DCH and AQRE demonstrates how hierarchical thinking models can generate statistically comparable predictions as equilibrium-based models.

## A.2 Likelihood Functions

This section derives the likelihood functions. For the cognitive hierarchy theories, I follow Camerer et al. (2004) to assume the prior distribution of levels follows Poisson distribution. Therefore, for both of the Poisson-DCH and the standard Poisson-CH, there is one parameter to be estimated—the average number of levels $\tau$. For AQRE, I follow McKelvey and Palfrey (1998) to estimate the logit-AQRE which has a single parameter $\lambda$.

### Poisson-CH Models

The Poisson-CH models assume each player's level is i.i.d. drawn from $(p_k)_{k=0}^{\infty}$ where

$$p_k \equiv \frac{e^{-\tau}\tau^k}{k!}, \quad \text{for all} \;\; k = 0, 1, 2, \dots$$

and $\tau > 0$. Because $\tau$ is the mean and variance of the Poisson distribution, the economic meaning of $\tau$ is the average level of sophistication among the population.

I first construct the likelihood function for the Poisson-DCH model. For each subject $i$, let $\Pi_i$ denote the set of information sets that subject $i$ has encountered in the game, and let $\mathcal{I}_i = (t, x_{-i})$ denote a generic information set. At any information set $\mathcal{I}_i$, subject $i$ can choose $c_i \in \{C, W\}$. Let $P_k(c_i|\mathcal{I}_i, \tau)$ be the probability of level $k$ players choosing $c_i$ at information set $\mathcal{I}_i$. Moreover, let $f(k|\mathcal{I}_i, \tau)$ be the posterior distribution of levels at information set $\mathcal{I}_i$. At period 1, $f(k|\mathcal{I}_i, \tau) = e^{-\tau}\tau^k/k!$. For later periods, $f(k|\mathcal{I}_i, \tau)$ given any $\tau$ can be analytically solved by Proposition 3 (two-person games) and Proposition X (three-person games). Finally, the predicted choice probability for $c_i$ at information set $\mathcal{I}_i$ is simply the aggregation of best responses from all levels weighted by the proportion $f(k|\mathcal{I}_i, \tau)$:

$$\mathcal{D}(c_i|\mathcal{I}_i, \tau) \equiv \sum_{k=0}^{\infty} f(k|\mathcal{I}_i, \tau) P_k(c_i|\mathcal{I}_i, \tau).$$

Consequently, the log-likelihood function for the DCH model can be formed by aggregating over every subject $i$, actions $c_i$ and information set $\mathcal{I}_i$:

$$\ln L^D(\tau) = \sum_i \sum_{\mathcal{I}_i \in \Pi_i} \sum_{c_i \in \{W,C\}} \mathbb{1}\{c_i, \mathcal{I}_i\} \ln \left[\mathcal{D}(c_i|\mathcal{I}_i, \tau)\right],$$

where $\mathbb{1}\{c_i, \mathcal{I}_i\}$ is the indicator function which is 1 when subject $i$ chooses $c_i$ at $\mathcal{I}_i$.

Second, the log-likelihood function for the standard Poisson-CH model can be constructed in the similar way. Given any $\tau$, the standard Poisson-CH model predicts a probability distribution over $\{1, \ldots, T, T+1\}$ (earliest period to choose $C$ or always $W$) for each level of players conditional on the announcement and other players' faces. Following previous notations, the probability of level $k$ subject $i$ choosing $t$ conditional on $x_{-i}$ is denoted by $\tilde{\sigma}_i^k(t|x_{-i})$, which can be analytically solved by Proposition 4 (two-person games) and Proposition Y (three-person games). Therefore, subject $i$'s predicted choice probability for $t \in \{1, \ldots, T, T+1\}$ conditional on $x_{-i}$ is the aggregation of choice frequencies of all levels weighted by Poisson($\tau$):

$$\tilde{\mathcal{S}}(t|x_{-i}, \tau) = \sum_{k=0}^{\infty} \frac{e^{-\tau}\tau^k}{k!} \tilde{\sigma}_i^k(t|x_{-i}).$$

Since $\tilde{\sigma}_i^0(t|x_{-i}) = \frac{1}{T+1}$ for all $t$, $\tilde{\mathcal{S}}(t|x_{-i}, \tau) > 0$ for all $t$. Moreover, the conditional probability to choose $C$ or $W$ at information set $\mathcal{I}_i = (t, x_{-i})$ can be computed by:

$$\mathcal{S}(C|\mathcal{I}_i, \tau) = \frac{\tilde{\mathcal{S}}(t|x_{-i}, \tau)}{\sum_{t' \geq t} \tilde{\mathcal{S}}(t'|x_{-i}, \tau)} \quad \text{and} \quad \mathcal{S}(W|\mathcal{I}_i, \tau) = 1 - \mathcal{S}(C|\mathcal{I}_i, \tau).$$

Finally, the log-likelihood function for the standard CH model can be constructed by aggregating over every subjects $i$, actions $c_i$, and information set $\mathcal{I}_i$:

$$\ln L^S(\tau) = \sum_i \sum_{\mathcal{I}_i \in \Pi_i} \sum_{c_i \in \{W,C\}} \mathbb{1}\{c_i, \mathcal{I}_i\} \ln \left[\mathcal{S}(c_i|\mathcal{I}_i, \tau)\right].$$

**Logit-AQRE Model**

For the purpose of illustrate, I only derive the likelihood function for two-person games. The likelihood function for three-person games can be derived by a similar calculation.

Let $Q(c_i|\mathcal{I}_i, \lambda)$ be the probability of subject $i$ choosing $c_i$ at information set $\mathcal{I}_i$ predicted by the logit-AQRE. In the two-person two-period dirty-faces game, each player's strategy is defined by a four-tuple $(q_1, q_2, r_1, r_2)$ which corresponds to $Q(C|1, O, \lambda)$, $Q(C|2, O, \lambda)$, $Q(C|1, X, \lambda)$, and $Q(C|2, X, \lambda)$, respectively. At information set $(t, x_{-i}) = (1, O)$, players would estimate the payoff of $C$ and $W$ by

$$U_{1,O}(C) = \alpha + \epsilon_{1,O,C}$$
$$U_{1,O}(W) = \delta\alpha(1 - r_1)q_2 + \epsilon_{1,O,U},$$

where $\epsilon_{1,O,C}$ and $\epsilon_{1,O,W}$ are independent random variables with a Weibull distribution with the precision parameter $\lambda$. Then the logit formula suggests

$$q_1 = \frac{1}{1 + exp\left\{\lambda\left[\delta\alpha(1 - r_1)q_2 - \alpha\right]\right\}}.$$

Similarly, $q_2$ can be expressed by:

$$q_2 = \frac{1}{1 + exp\left\{-\delta\alpha\lambda\right\}}.$$

On the other hand, when observing a dirty face and the game proceeds to period 2, players' posterior beliefs become:

$$\mu \equiv \Pr(X|2, X) = \frac{p(1 - r_1)}{p(1 - r_1) + (1 - p)(1 - q_1)} = \frac{1}{1 + \left(\frac{1-p}{p}\right)\left(\frac{1-q_1}{1-r_1}\right)},$$

and hence the expected payoff to choose $C$ at information set $(2, X)$ is:

$$\delta\left[\alpha\mu - (1 - \mu)\right] = \delta\left[(1 + \alpha)\mu - 1\right].$$

As a result, $r_2$ satisfies that

$$r_2 = \frac{1}{1 + exp\left\{\lambda\delta\left[1 - (1 + \alpha)\mu\right]\right\}}.$$

Finally, the expected payoff of choosing $C$ at information set $(1, X)$ is $\alpha p - (1 - p)$, while the expected payoff of $W$ is

$$\underbrace{\left[p(1 - r_1) + (1 - p)(1 - q_1)\right]}_{\text{prob. to reach period 2}} r_2\delta\left[(1 + \alpha)\mu - 1\right] \equiv A,$$

and therefore, $r_1$ can be expressed by:

$$r_1 = \frac{1}{1 + exp\left\{\lambda\left[A + (1 - p) - \alpha p\right]\right\}}.$$

As plugging $p = 2/3$, $\delta = 4/5$ and $\alpha = 2/3$ into the choice probabilities, we can obtain that

$$\begin{aligned} r_1 &= \frac{1}{1 + exp\left\{\lambda\left[\frac{2}{15}(1-r_1)r_2 - \frac{4}{15}(1-q_1)r_2 + \frac{1}{6}\right]\right\}} \\ r_2 &= \frac{1}{1 + exp\left\{\lambda\left[\frac{4}{5} - \frac{2-2r_1}{3-2r_1-q_1}\right]\right\}} \\ q_1 &= \frac{1}{1 + exp\left\{\lambda\left[\frac{1}{5}(1-r_1)q_2 - \frac{1}{4}\right]\right\}} \\ q_2 &= \frac{1}{1 + exp\left\{-\frac{1}{5}\lambda\right\}}. \end{aligned}$$

Given each $\lambda$, the system of four equations with four unknowns can be solved uniquely. Besides, for each $\mathcal{I}_i$, $Q(W|\mathcal{I}_i, \lambda) = 1 - Q(C|\mathcal{I}_i, \lambda)$. Thus, the log-likelihood function can be formed by aggregating over every subject $i$, action $c_i$, and information set $\mathcal{I}_i$:

$$\ln L^Q(\lambda) = \sum_i \sum_{\mathcal{I}_i \in \Pi_i} \sum_{c_i \in \{W,C\}} \mathbb{1}\{c_i, \mathcal{I}_i\} \ln\left[Q(c_i|\mathcal{I}_i, \lambda)\right].$$

## A.3 Estimation Results

The Poisson-DCH, standard Poisson-CH and the AQRE models are estimated by maximum likelihood estimation. Table A.2 reports the estimation results on Treatment 1 and Treatment 2 data, showing the estimated parameters and the fitness of each model. Comparing the fitness of these models, I find that the log-likelihood of DCH is significantly higher than standard CH (Vuong Test p-value $< 0.001$ for both treatments), while it is not significantly different from AQRE (Treatment 1: p-value $= 0.144$; Treatment 2: p-value $= 0.184$). This result suggests in both Treatment 1 and 2, DCH outperforms the standard CH in capturing the empirical patterns and generates predictions that are statistically comparable to other equilibrium-based behavioral solution concepts.

Comparing the estimation results of Treatment 1 and 2, I observe that there is more randomness in three-person games compared to two-person games. In two-person games, the DCH estimates indicate that players can think 1.269 steps (95% C.I. $= [1.093, 1.445]$) on average, while in three-person games, players can only think an average of 0.370 steps (95% C.I. $= [0.286, 0.454]$). Additionally, the estimation result of AQRE suggests that as the game changes from two-person games to three-person games, the precision of decision-making decreases significantly (from 7.663 to 5.278). This implies that players are less likely to make best responses in three-person games.

To analyze the differences between the models in detail, I compare the choice probabil-

Table A.2: Estimation Results for Treatment 1 and Treatment 2 Data

| | | Two-Person Games | | | Three-Person Games | | |
|---|---|---|---|---|---|---|---|
| | | DCH | Standard CH | AQRE | DCH | Standard CH | AQRE |
| Parameters | $\tau$ | 1.269 | 1.161 | — | 0.370 | 0.140 | — |
| | S.E. | (0.090) | (0.095) | — | (0.043) | (0.039) | — |
| | $\lambda$ | — | — | 7.663 | — | — | 5.278 |
| | S.E. | — | — | (0.493) | — | — | (0.404) |
| Fitness | LL | -360.75 | -381.46 | -368.38 | -575.30 | -608.45 | -565.05 |
| | AIC | 723.50 | 764.91 | 738.76 | 1152.61 | 1218.89 | 1132.11 |
| | BIC | 728.04 | 769.45 | 743.29 | 1157.43 | 1223.72 | 1136.93 |
| Vuong Test | | | 6.517 | 1.463 | | 3.535 | -1.330 |
| p-value | | | $< 0.001$ | 0.144 | | $< 0.001$ | 0.184 |

Note: There are 294 games (rounds × groups) in Treatment 1 and 224 games in Treatment 2.

ities predicted by each model. Figure A.2 illustrates the choice probabilities in two-person games, while Figure A.3 displays the choice probabilities in three-person games. Comparing the DCH and the standard CH models, I observe that the standard CH model generally underestimates the probability of choosing $C$ in period 1. In two-person games, the empirical frequencies of choosing $C$ at information sets $(1, O)$ and $(1, X)$ are 0.943 and 0.210, respectively. Yet, the predictions of the standard CH model are 0.791 and 0.104 for the same information sets. A similar pattern of underestimation is also evident in three-person games.

The underestimation is primarily caused by the difference in the specifications of level 0 players' behavior. In two-person games, the standard CH model assumes that level 0 players uniformly randomize across the set $\{1, 2, 3\}$. Consequently, the probability of level 0 players choosing $C$ in period 1 according to the standard CH model is $1/3$. In contrast, in the DCH model, level 0 players uniformly randomize at every information set, resulting in a probability of $1/2$ for them to choose $C$. Similarly, in three-person games, the standard CH model assumes that level 0 players uniformly randomize across the set $\{1, 2, 3, 4\}$, leading to a probability of $1/4$ for them to choose $C$ in period 1. In contrast, in the DCH model, level 0 players' behavior remains the same across both two-person and three-person games. These differences in level 0 players' behavior contribute to the underestimation of the probability of choosing $C$ in the standard CH model compared to DCH.

Moreover, the key difference between the CH approach and AQRE is highlighted in the off-equilibrium-path information sets.[33] Conceptually, the reason why the game could

[33] When $x_{-i} = OO$, the equilibrium predicts that players will choose $C$ in period 1, resulting in the game not proceeding beyond period 2. Similarly, when $x_{-i} = OX$, the equilibrium suggests that players should choose $C$ in period 2, preventing the game from progressing to period 3.

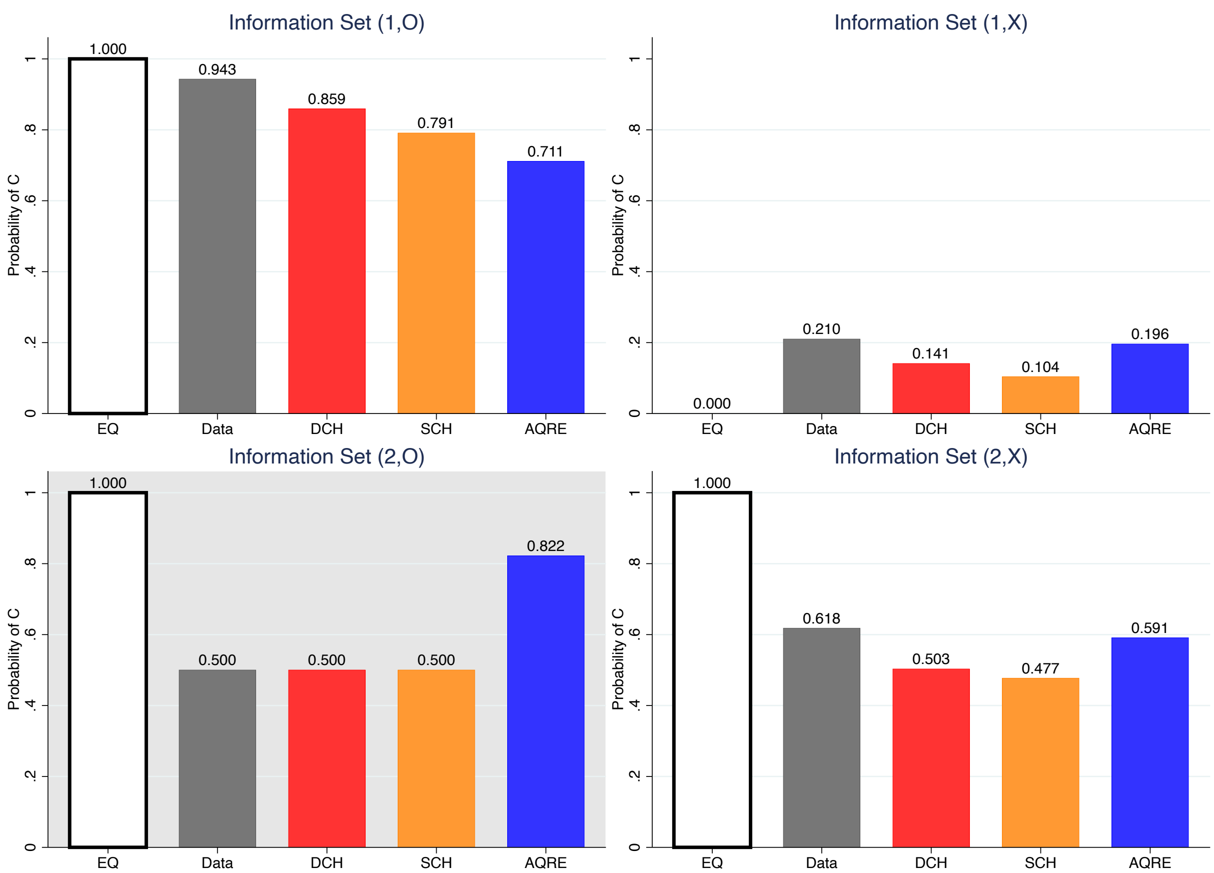



Figure A.2: The choice probabilities in two-person games at different information sets. Each panel plots the empirical choice frequencies and the predictions of different models at one information set. The gray panel represents the off-equilibrium path information set.

proceed to the off-equilibrium-path information sets differs between the CH approach and AQRE. From the perspective of AQRE, the off-equilibrium-path information sets are reached due to mistakes. As a result, AQRE predicts a high probability of choosing $C$ at these off-equilibrium-path information sets because the expected payoff of choosing $C$ is much higher than $W$ at these information sets. By contrary, in the CH approach, the off-equilibrium-path information sets are reached because the players are not sophisticated enough. For instance, when observing no dirty face, players should immediately choose $C$ since it is a dominant strategy. If someone doesn't choose $C$, they are definitely a level 0 player.

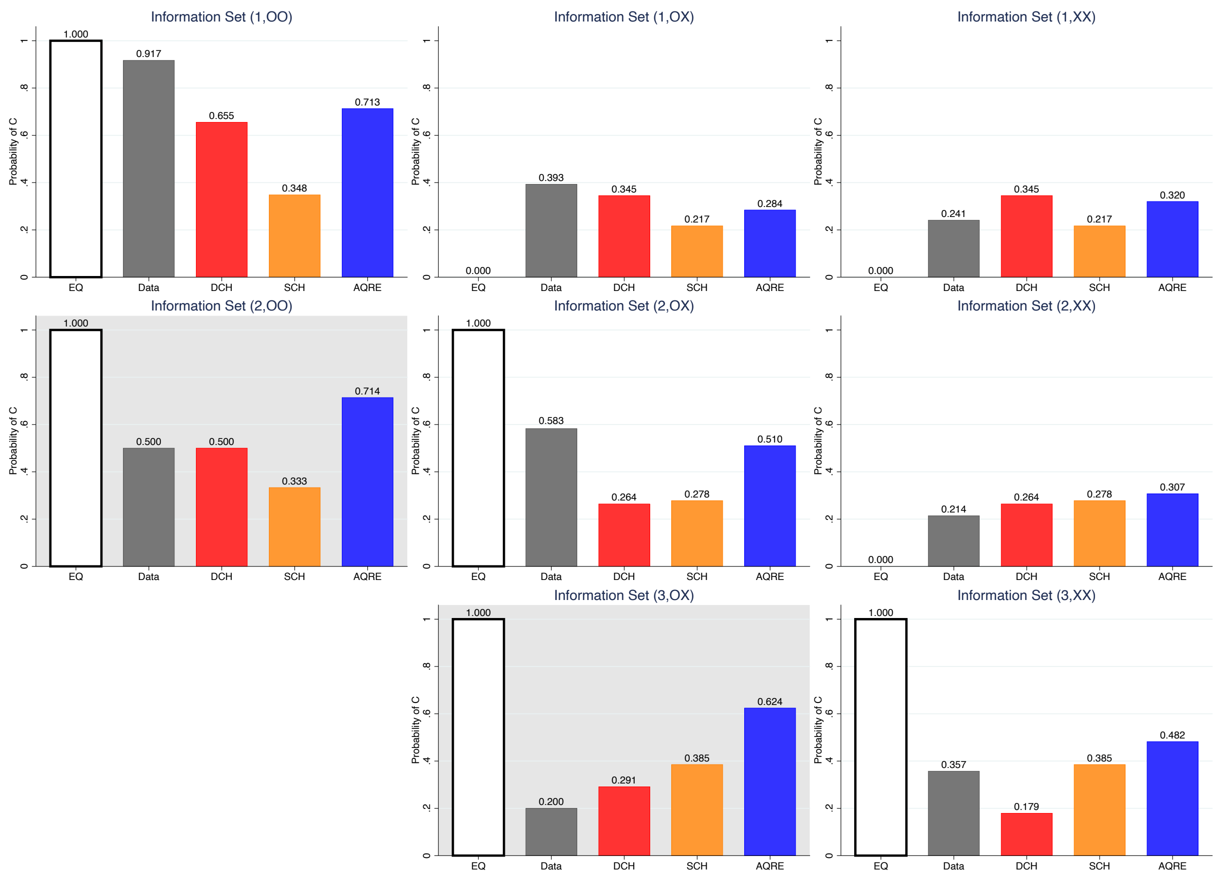



Figure A.3: The choice probabilities in three-person games at different information sets. Each panel plots the empirical choice frequencies and the predictions of different models at one information set. The gray panels represent the off-equilibrium-path information sets.

From the choice probabilities, it can be observed that DCH provides the most accurate predictions at off-path information sets, regardless of whether it is in two-person or three-person games. At information sets $(2, O)$ and $(2, OO)$, the empirical choice probabilities of $C$ are 0.5, which are correctly predicted by DCH. Furthermore, at the information set $(3, OX)$, the empirical choice probability of $C$ is 0.2, while the predictions of DCH, standard CH, and

AQRE are 0.291, 0.385, and 0.624, respectively.

Table A.3: Estimation Results for Pooled Data

| | | DCH | Standard CH | AQRE |
|---|---|---|---|---|
| Parameters | $\tau$ | 1.030 | 0.241 | — |
| | S.E. | (0.060) | (0.033) | — |
| | $\lambda$ | — | — | 6.235 |
| | S.E. | — | — | (0.302) |
| Fitness | LL | -956.92 | -1047.12 | -940.65 |
| | AIC | 1915.84 | 2096.23 | 1883.30 |
| | BIC | 1921.22 | 2101.62 | 1888.69 |
| Vuong Test | | | 7.513 | -1.363 |
| p-value | | | $< 0.001$ | 0.173 |
| LR Test $\chi^2_{(1)}$ | | 41.74 | 114.42 | 14.44 |
| p-value | | $< 0.001$ | $< 0.001$ | $< 0.001$ |

Note: The likelihood ratio test is testing if the log-likelihood of two-parameter models (Treatment 1 and 2) is significantly higher than the log-likelihood of one-parameter models.

In addition, I estimate the three models using the pooled data, and the results are reported in Table A.3. Consistent with the results from the two-person games and three-person games, it can be observed that DCH provides a significantly better fit to the data compared to the standard CH model (Vuong test: p-value $< 0.001$). However, there is no statistically significant difference between DCH and AQRE (Vuong test p-value $= 0.173$). Furthermore, I conduct a likelihood ratio test on all three models to assess whether allowing different parameters for two-person and three-person games can significantly improve the model fit. The results indicate that the heterogeneous models are significantly better than the homogeneous models. Taken together, these findings lead to the conclusion that both the level of sophistication and the precision vary with the complexity of the games.

To summarize, it is not surprising that DCH can provide a better explanation for the data compared to the misspecified standard CH model in dynamic games. However, what is surprising is that when the CH model is correctly specified, the estimated average level of sophistication is 1.03, which falls within the expected range of a "regular" $\tau$ value between 1 and 2, as predicted by Camerer et al. (2004).

# B Supplementary Analysis for Experimental Data (For Online Publication)

## B.1 Supplementary Tables

## B.2 Likelihood Functions

### Quantal Cursed Sequential Equilibrium

The "Quantal Cursed Sequential Equilibrium (QCSE)" is a model applicable to multi-stage games with observed actions. This model relaxes both the requirements of best responses and Bayesian inference. Specifically, QCSE is a hybrid model, combining the Agent Quantal Response Equilibrium (AQRE) proposed by McKelvey and Palfrey (1998) and the Cursed Sequential Equilibrium introduced by Fong et al. (2025).

Consider an assessment $(\mu, \sigma)$. For any player $i$ and any history $h^{t-1}$, the *average behavioral strategy profile of $-i$* is defined as:

$$\bar{\sigma}_{-i}(a^t_{-i}|\theta_i, h^{t-1}) = \sum_{\theta_{-i}\in\Theta_{-i}} \mu_i(\theta_{-i}|\theta_i, h^{t-1})\sigma_{-i}(a^t_{-i}|\theta_{-i}, h^{t-1}).$$

In QCSE, players have incorrect perceptions about the behavioral strategies of other players. Instead of thinking they are using $\sigma_{-i}$, a $\chi$-cursed type $\theta_i$ player $i$ would believe the other players are using a $\chi$-weighted average of the average behavioral strategy and the true behavioral strategy:

$$\sigma^{\chi}_{-i}(a^t_{-i}|\theta_{-i}, \theta_i, h^{t-1}) = \chi\bar{\sigma}_{-i}(a^t_{-i}|\theta_i, h^{t-1}) + (1-\chi)\sigma_{-i}(a^t_{-i}|\theta_{-i}, h^{t-1}).$$

The beliefs of player $i$ about $\theta_{-i}$ in QCSE are updated via Bayes' rule, whenever possible, assuming other players are using the $\chi$-cursed behavioral strategy rather than the true behavioral strategy. This updating rule is called the *$\chi$-cursed Bayes' rule*. Specifically, an assessment satisfies the $\chi$-cursed Bayes' rule if the belief system is derived from the Bayes' rule while perceiving others are using $\sigma^{\chi}_{-i}$ rather than $\sigma_{-i}$.

Consider any totally mixed strategy profile $\sigma \in \Sigma^0$. As shown by Fong et al. (2025), if the belief system $\mu$ is derived from the *$\chi$-cursed Bayes' rule*, then player $i$'s cursed belief is simply a linear combination of player $i$'s cursed belief at the beginning of that stage (with $\chi$

weight) and the Bayesian posterior belief (with $1-\chi$ weight). That is, for any $h^t = (h^{t-1}, a^t)$,

$$\mu_i(\theta_{-i}|\theta_i, h^t) = \chi\mu_i(\theta_{-i}|\theta_i, h^{t-1}) + (1-\chi)\left[\frac{\mu_i(\theta_{-i}|\theta_i, h^{t-1})\sigma_{-i}(a^t_{-i}|\theta_{-i}, h^{t-1})}{\sum_{\theta'_{-i}}\mu_i(\theta'_{-i}|\theta_i, h^{t-1})\sigma_{-i}(a^t_{-i}|\theta'_{-i}, h^{t-1})}\right].$$

For any player $i$, any $\chi \in [0,1]$, $\sigma \in \Sigma^0$, and type profile $\theta \in \Theta$, let $\rho_i^\chi(h^T|\theta, h^t, \sigma^\chi_{-i}, \sigma_i)$ be $i$'s perceived conditional realization probability of terminal history $h^T \in \mathcal{H}^T$ at history $h^t \in \mathcal{H}\backslash\mathcal{H}^T$ if the type profile is $\theta$ and $i$ uses the behavioral strategy $\sigma_i$ whereas perceives other players' using the cursed behavioral strategy $\sigma^\chi_{-i}$. At every non-terminal history $h^t$, a $\chi$-cursed player in QCSE will use $\chi$-cursed Bayes' rule to derive the posterior belief about the other players' types. Accordingly, a type $\theta_i$ player $i$'s conditional expected payoff at history $h^t$ is:

$$\bar{u}_i(\sigma|\theta_i, h^t) \equiv \sum_{\theta_{-i}\in\Theta_{-i}}\sum_{h^T\in\mathcal{H}^T}\mu_i(\theta_{-i}|\theta_i, h^t)\rho_i^\chi(h^T|\theta, h^t, \sigma^\chi_{-i}, \sigma_i)u_i(h^T, \theta_i, \theta_{-i}).$$

Moreover, let $\bar{u}_i(a, \sigma|\theta_i, h^t)$ be the conditional expected payoff of player $i$ of using $a \in A_i(h^t)$ with probability one, and using $\sigma_i$ elsewhere.

In QCSE, there is a parameter $\lambda \in [0,\infty)$ that governs the precision of choices. Given an assessment $(\mu, \sigma)$ where $\sigma \in \Sigma^0$ and $\mu$ is derived from the $\chi$-cursed Bayes' rule, if $(\mu, \sigma)$ is a QCSE for any player $i$, history $h^t$, and type $\theta_i$, then type $\theta_i$ player $i$ will have choice probabilities at $h^t$ that follow a multinomial logit distribution. In particular, the probability of player $i$ choosing $a \in A_i(h^t)$ is

$$\frac{e^{\lambda\bar{u}_i(a,\sigma|\theta_i,h^t)}}{\sum_{a'\in A_i(h^t)}e^{\lambda\bar{u}_i(a',\sigma|\theta_i,h^t)}}.$$

In summary, for each $\lambda \in [0,\infty)$ and $\chi \in [0,1]$, an assessment $(\mu, \sigma)$ is a QCSE if

1. The belief system is derived from the $\chi$-cursed Bayes' rule, and

2. For any player $i$, type $\theta_i$, history $h^t$ and $a \in A_i(h^t)$,

$$\sigma_i(a|\theta_i, h^t) = \frac{e^{\lambda\bar{u}_i(a,\sigma|\theta_i,h^t)}}{\sum_{a'\in A_i(h^t)}e^{\lambda\bar{u}_i(a',\sigma|\theta_i,h^t)}}.$$

When estimating QCSE, constructing the likelihood function follows a similar process as described in Appendix A.2. For each information set $\mathcal{I}_i$, QCSE uniquely predicts the choice probability of each $a_i$, denoted as $\bar{Q}(a_i|\mathcal{I}_i, \lambda, \chi)$, given $\lambda$ and $\chi$. The log-likelihood function can be formed by aggregating over every subject $i$, action $a_i$, and information set $\mathcal{I}_i$:

$$\ln L^{\bar{Q}}(\lambda, \chi) = \sum_i \sum_{\mathcal{I}_i \in \Pi_i} \sum_{a_i \in A_i(\mathcal{I}_i)} \mathbb{1}\{a_i, \mathcal{I}_i\} \ln \left[\bar{Q}(a_i|\mathcal{I}_i, \lambda, \chi)\right].$$

**Quantal Dynamic Cognitive Hierarchy Solution**

The "Quantal Dynamic Cognitive Hierarchy Solution (QDCH)" is a natural extension of DCH, where all strategic levels of players make quantal responses instead of best responses. In particular, following previous notations, for any $i \in N$, $\tau_i \geq 1$, $\theta \in \Theta$, $\sigma$, and $\tau_{-i}$ such that $\tau_j < \tau_i$ for any $j \neq i$, let $P_i^{\tau_i}(h^T|\theta, h^{t-1}, \tau_{-i}, \sigma_{-i}^{-\tau_i}, \sigma_i^{\tau_i})$ be level $\tau_i$ player $i$'s belief about the conditional realization probability of $h^T \in \mathcal{H}^T$ at history $h^{t-1} \in \mathcal{H} \backslash \mathcal{H}^T$ if the type profile is $\theta$, the level profile is $\tau$, and player $i$ uses $\sigma_i^{\tau_i}$. In this case, level $\tau_i$ player $i$'s expected payoff at any $h^t \in \mathcal{H} \backslash \mathcal{H}^T$ is:

$$\bar{u}_i^{\tau_i}(\sigma|\theta_i, h^t) \equiv$$
$$\sum_{h^T \in \mathcal{H}^T} \sum_{\theta_{-i} \in \Theta_{-i}} \sum_{\{\tau_{-i}: \tau_j < k \ \forall j \neq i\}} \mu_i^{\tau_i}(\theta_{-i}, \tau_{-i} \mid \theta_i, h^t) P_i^{\tau_i}(h^T|\theta, h^t, \tau_{-i}, \sigma_{-i}^{-\tau_i}, \sigma_i^{\tau_i}) u_i(h^T, \theta_i, \theta_{-i}).$$

Similar to QCSE, in QDCH, there is a parameter $\lambda \in [0, \infty)$ that governs the precision of choices. Let $\bar{u}_i^{\tau_i}(a, \sigma|\theta_i, h^t)$ be the conditional expected payoff of level $\tau_i$ player $i$ of using $a \in A_i(h^t)$ with probability one, and using $\sigma_i^{\tau_i}$ elsewhere. In QDCH, players' choice probabilities follow multinomial logit distributions. That is, in QDCH, the probability of level $\tau_i$ player $i$ choosing $a \in A_i(h^t)$ at history $h^t$ is

$$\sigma_i^{\tau_i}(a|\theta_i, h^t) = \frac{e^{\lambda \bar{u}_i^{\tau_i}(a, \sigma|\theta_i, h^t)}}{\sum_{a' \in A_i(h^t)} e^{\lambda \bar{u}_i^{\tau_i}(a', \sigma|\theta_i, h^t)}}.$$

When estimating QDCH, I assume the prior distribution of levels follows Poisson($\tau$). At any information set $\mathcal{I}_i$, let $f(k|\mathcal{I}_i, \lambda, \tau)$ be the posterior distribution of levels at information set $\mathcal{I}_i$ given $\lambda$ and $\tau$. In this case, the predicted choice probability for $a_i$ at $\mathcal{I}_i$ is the aggregation of quantal responses from all levels weighted by the proportion $f(k|\mathcal{I}_i, \lambda, \tau)$:

$$\bar{\mathcal{D}}(a_i|\mathcal{I}_i, \lambda, \tau) \equiv \sum_{k=0}^{\infty} f(k|\mathcal{I}_i, \lambda, \tau) P_k(a_i|\mathcal{I}_i, \lambda, \tau),$$

where $P_k(a_i|\mathcal{I}_i, \lambda, \tau)$ is the probability of level $k$ players choosing $a_i$ at $\mathcal{I}_i$. Consequently, the log-likelihood function for QDCH can be formed by aggregating over every subject $i$, actions

# Supplemental Appendix for "Cognitive Hierarchies in Multi-Stage Games of Incomplete Information: Theory and Experiment"

Po-Hsuan Lin*

October 22, 2025

## Contents

1 Omitted Proofs 1

2 Additional Analysis 11

2.1 Additional Result for Poisson-DCH . . . 11

2.2 Supplementary Tables . . . 13

3 Experimental Instructions 15

3.1 Sequential Treatment . . . 15

3.2 Simultaneous Treatment . . . 18

3.3 Screenshots . . . 21

# 1 Omitted Proofs

## Proof of Proposition 3

***Step 1:*** Consider any $i \in N$. If $x_{-i} = O$, then player $i$ knows his face is dirty immediately. Therefore, $C$ is a dominant strategy, suggesting $\sigma_i^k(O, t) = 1$ for all $k \geq 1$ and $1 \leq t \leq T$. If $x_{-i} = X$, player $i$'s belief of having a dirty face at period 1 is $d$. Hence, the expected payoff of choosing $C$ at period 1 is $d\alpha - (1 - d) < 0$, implying $\sigma_i^k(X, 1) = 0$ for all $k \geq 1$. Finally, since level-1 players believe the other player's actions don't convey any information about their own face types, the expected payoff of $C$ at each period is $d\alpha - (1 - d) < 0$, implying $\sigma_i^1(X, t) = 0$ for any $1 \leq t \leq T$.

*Department of Economics, University of Virginia, Charlottesville, VA 22904. Email: plin@virginia.edu

***Step 2:*** Consider any level-$k \geq 2$, and period $2 \leq t \leq T$. This step characterizes the DCH posterior belief when $x_{-i} = X$. When the game proceeds to period $t$, the posterior belief of $(x_i, \tau_{-i}) = (f, l)$ for any $f \in \{O, X\}$ and $0 \leq l \leq k-1$ is:

$$\mu^k(f, l|X, t) = \frac{\left[\prod_{t'=1}^{t-1}(1 - \sigma_{-i}^l(f, t'))\right] p_l \Pr(f)}{\sum_{x \in \{O,X\}} \sum_{j=0}^{k-1} \left[\prod_{t'=1}^{t-1}(1 - \sigma_{-i}^j(x, t'))\right] p_j \Pr(x)}. \tag{O.1}$$

By step 1, since strategic players will claim immediately when seeing a clean face, $\sigma_{-i}^l(O, t') = 1$ for all $1 \leq t' \leq t-1$. Therefore, as the game proceeds to period $t \geq 2$, level-$k$ player $i$ would know that it is impossible for the other player to observe a dirty face and have a positive level of sophistication at the same time. Furthermore, let $\mathcal{M}_i^k(t)$ be the support of level-$k$ player $i$'s marginal belief of $\tau_{-i}$ at period $t$. For any $0 \leq l \leq k-1$,

$$l \in \mathcal{M}_i^k(t) \iff \sum_{x_i \in \{O,X\}} \prod_{t'=1}^{t-1}(1 - \sigma_{-i}^l(x_i, t')) > 0,$$

and we let $\mathcal{M}_{i+}^k(t) \equiv \mathcal{M}_i^k(t) \backslash \{0\}$. Therefore, equation (O.1) implies that for any $t \geq 2$,

$$\mu^k(X, 0|X, t) = \frac{\left(\frac{1}{2}\right)^{t-1} dp_0}{\left(\frac{1}{2}\right)^{t-1} p_0 + d\sum_{j \in \mathcal{M}_{i+}^k(t)} p_j}, \quad \mu^k(O, 0|X, t) = \frac{\left(\frac{1}{2}\right)^{t-1}(1-d)p_0}{\left(\frac{1}{2}\right)^{t-1} p_0 + d\sum_{\in \mathcal{M}_{i+}^k(t)} p_j}.$$

Moreover, for any $1 \leq k' \leq k-1$, $\mu^k(O, k'|X, t) = 0$, and for any $l \in \mathcal{M}_{i+}^k(t)$,

$$\mu^k(X, l|X, t) = \frac{dp_l}{\left(\frac{1}{2}\right)^{t-1} p_0 + d\sum_{j \in \mathcal{M}_{i+}^k(t)} p_j}.$$

Consequently, the marginal belief of having a dirty face at period $2 \leq t \leq T$ is:

$$\mu^k(X|X, t) = \sum_{j=0}^{k-1} \mu^k(X, j|X, t) = \frac{d\left[\left(\frac{1}{2}\right)^{t-1} p_0 + \sum_{j \in \mathcal{M}_{i+}^k(t)} p_j\right]}{\left(\frac{1}{2}\right)^{t-1} p_0 + d\sum_{j \in \mathcal{M}_{i+}^k(t)} p_j}.$$

Thus, the expected payoff of choosing $C$ at period $t$ is $\delta^{t-1}\left[(1+\alpha)\mu^k(X|X, t) - 1\right]$, which equals to $\mathbb{E}u_i^k(C|X, t) =$

$$\frac{\delta^{t-1}}{\left(\frac{1}{2}\right)^{t-1} p_0 + d\sum_{j \in \mathcal{M}_{i+}^k(t)} p_j}\left\{d\alpha\left[\left(\frac{1}{2}\right)^{t-1} p_0 + \sum_{j \in \mathcal{M}_{i+}^k(t)} p_j\right] - (1-d)\left[\left(\frac{1}{2}\right)^{t-1} p_0\right]\right\}. \tag{O.2}$$

Finally, at period $t$, level-$k$ player $i$ believes the other player will wait with probability

$$\frac{1}{2}\mu^k(0|X, t) + \sum_{j \in \mathcal{M}_{i+}^k(t+1)} \mu^k(j|X, t) = \frac{\left(\frac{1}{2}\right)^t p_0 + d\sum_{j \in \mathcal{M}_{i+}^k(t+1)} p_j}{\left(\frac{1}{2}\right)^{t-1} p_0 + d\sum_{j \in \mathcal{M}_{i+}^k(t)} p_j}. \tag{O.3}$$

***Step 3:*** This step proves a monotonicity result: if $\sigma_i^k(X,t)=1$, then $\sigma_i^{k+1}(X,t)=1$ for any $k \geq 2$ and $2 \leq t \leq T$. The proof consists of two cases. First consider period $T$. Equation (O.2) implies $\sigma_i^k(X,T)=1$ if and only if

$$\frac{\delta^{T-1}}{\left(\frac{1}{2}\right)^{T-1} p_0 + d\sum_{j\in\mathcal{M}_{i+}^k(T)} p_j}\left\{d\alpha\left[\left(\frac{1}{2}\right)^{T-1} p_0 + \sum_{j\in\mathcal{M}_{i+}^k(T)} p_j\right] - (1-d)\left[\left(\frac{1}{2}\right)^{T-1} p_0\right]\right\} \geq 0$$

$$\iff \bar{\alpha} \;\geq\; \frac{\left(\frac{1}{2}\right)^{T-1} p_0}{\left(\frac{1}{2}\right)^{T-1} p_0 + \sum_{j\in\mathcal{M}_{i+}^k(T)} p_j}.$$

Because $\mathcal{M}_i^k(T) \subseteq \mathcal{M}_i^{k+1}(T)$, it can be proven that

$$\bar{\alpha} \;\geq\; \frac{\left(\frac{1}{2}\right)^{T-1} p_0}{\left(\frac{1}{2}\right)^{T-1} p_0 + \sum_{j\in\mathcal{M}_{i+}^k(T)} p_j} \;\geq\; \frac{\left(\frac{1}{2}\right)^{T-1} p_0}{\left(\frac{1}{2}\right)^{T-1} p_0 + \sum_{j\in\mathcal{M}_{i+}^{k+1}(T)} p_j},$$

implying that if it is optimal for level-$k$ player $i$ to claim at period $T$, it is also optimal for level-$(k+1)$ player $i$ to claim.

Second, consider any period $2 \leq t \leq T-1$. Note that if level-$k$ players would choose $C$ at period $t$, $k \notin \mathcal{M}_i^{k+1}(t+1)$ and hence $\mathcal{M}_{i+}^k(t') = \mathcal{M}_{i+}^{k+1}(t')$ for any $t+1 \leq t' \leq T$. Therefore, as the game proceeds beyond period $t$, level-$k$ and level-$(k+1)$ players will have the same continuation value. Let $V_{\tilde{t}}^{\tilde{k}}$ be level-$\tilde{k}$ players' continuation value at period $\tilde{t}$. The observation implies $V_{t+1}^k = V_{t+1}^{k+1}$. Coupled with that $\mathcal{M}_{i+}^k(t) \subseteq \mathcal{M}_{i+}^{k+1}(t)$, level-$(k+1)$ player $i$'s expected payoff of $W$ at period $t$ satisfies

$$\frac{\left(\frac{1}{2}\right)^{t} p_0 + d\sum_{j\in\mathcal{M}_{i+}^{k+1}(t+1)} p_j}{\left(\frac{1}{2}\right)^{t-1} p_0 + d\sum_{j\in\mathcal{M}_{i+}^{k+1}(t)} p_j} V_{t+1}^{k+1} \leq \frac{\left(\frac{1}{2}\right)^{t} p_0 + d\sum_{j\in\mathcal{M}_{i+}^{k}(t+1)} p_j}{\left(\frac{1}{2}\right)^{t-1} p_0 + d\sum_{j\in\mathcal{M}_{i+}^{k}(t)} p_j} V_{t+1}^{k},$$

where the RHS is level-$k$ players' expected payoff of choosing $W$ at period $t$. The inequality shows level-$k$ players' expected payoff of choosing $W$ is weakly higher than level-$(k+1)$ players' expected payoff of choosing $W$. To complete the proof, it suffices to argue that level-$(k+1)$ players' expected payoff of $C$ at period $t$ is higher than level-$k$ players' expected payoff of $C$. This is true because $\mathcal{M}_{i+}^k(t) \subseteq \mathcal{M}_{i+}^{k+1}(t)$ implies $\mu^{k+1}(X|X,t) \geq \mu^k(X|X,t)$, and hence,

$$\delta^{t-1}\left[(1+\alpha)\mu^{k+1}(X|X,t) - 1\right] \geq \delta^{t-1}\left[(1+\alpha)\mu^{k}(X|X,t) - 1\right].$$

***Step 4:*** The proposition is proven by induction on $k$. This step establishes the base case for level-2 players. By step 1, $\sigma_i^1(X,t) = 0$ for all $1 \leq t \leq T$, and hence $\mathcal{M}_{i+}^2(t) = \{1\}$ for all $1 \leq t \leq T$. Therefore, equation (O.2) suggests the expected payoff of $C$ at period $T$ is

$$\mathbb{E}u_i^2(C|X,T) = \frac{\delta^{T-1}}{\left(\frac{1}{2}\right)^{T-1} p_0 + dp_1}\left\{d\alpha\left[\left(\frac{1}{2}\right)^{T-1} p_0 + p_1\right] - (1-d)\left[\left(\frac{1}{2}\right)^{T-1} p_0\right]\right\}.$$

Therefore, $C$ is optimal at period $T$ if and only if

$$\mathbb{E}u_i^2(C|X,T) \geq 0 \iff \bar{\alpha} \geq \frac{\left(\frac{1}{2}\right)^{T-1} p_0}{\left(\frac{1}{2}\right)^{T-1} p_0 + p_1}.$$

For any period $2 \leq t \leq T-1$, I first prove the direction of necessity. Equation (O.3) implies level-2 player $i$'s belief about the other player choosing $W$ at period $t$ is:

$$\frac{1}{2}\mu^2(0|X,t) + \mu^2(1|X,t) = \frac{\left(\frac{1}{2}\right)^{t} p_0 + dp_1}{\left(\frac{1}{2}\right)^{t-1} p_0 + dp_1}.$$

Therefore, the expected payoff of $W$ at period $t$ is *at least* $\left[\frac{\left(\frac{1}{2}\right)^{t} p_0 + dp_1}{\left(\frac{1}{2}\right)^{t-1} p_0 + dp_1}\right] \mathbb{E}u_i^2(C|X,t+1) =$

$$\frac{\delta^t}{\left(\frac{1}{2}\right)^{t-1} p_0 + dp_1} \left\{ d\alpha \left[ \left(\frac{1}{2}\right)^t p_0 + p_1 \right] - (1-d) \left[ \left(\frac{1}{2}\right)^t p_0 \right] \right\}.$$

Since $W$ is always available, $C$ is strictly dominated at period $t$ for level-2 player $i$ if

$$\mathbb{E}u_i^2(C|X,t) < \left[\frac{\left(\frac{1}{2}\right)^{t} p_0 + dp_1}{\left(\frac{1}{2}\right)^{t-1} p_0 + dp_1}\right] \mathbb{E}u_i^2(C|X,t+1) \iff \bar{\alpha} < \frac{\left[\left(\frac{1}{2}\right)^{t-1} - \left(\frac{1}{2}\right)^{t} \delta\right] p_0}{\left[\left(\frac{1}{2}\right)^{t-1} - \left(\frac{1}{2}\right)^{t} \delta\right] p_0 + (1-\delta)p_1}.$$

This proves the direction of necessity.

Second, the sufficiency is proven by induction on the periods. Namely, I show the sufficiency holds for any period $T-t'$ where $1 \leq t' \leq T-2$. Consider the base case for period $T-1$. Because

$$\bar{\alpha} \geq \frac{\left[\left(\frac{1}{2}\right)^{T-2} - \left(\frac{1}{2}\right)^{T-1} \delta\right] p_0}{\left[\left(\frac{1}{2}\right)^{T-2} - \left(\frac{1}{2}\right)^{T-1} \delta\right] p_0 + (1-\delta)p_1} > \frac{\left(\frac{1}{2}\right)^{T-1} p_0}{\left(\frac{1}{2}\right)^{T-1} p_0 + p_1},$$

level-2 players will choose $C$ at period $T$, so it is optimal to choose $C$ at period $T-1$ if

$$\mathbb{E}u_i^2(C|X,T-1) \geq \left[\frac{\left(\frac{1}{2}\right)^{T-1} p_0 + dp_1}{\left(\frac{1}{2}\right)^{T-2} p_0 + dp_1}\right] \mathbb{E}u_i^2(C|X,T) \iff \bar{\alpha} \geq \frac{\left[\left(\frac{1}{2}\right)^{T-2} - \left(\frac{1}{2}\right)^{T-1} \delta\right] p_0}{\left[\left(\frac{1}{2}\right)^{T-2} - \left(\frac{1}{2}\right)^{T-1} \delta\right] p_0 + (1-\delta)p_1}.$$

Now, suppose there is $t' \leq T-2$ such that the statement holds at any period $T-t$ where $1 \leq t \leq t'-1$. It can be proven that the sufficiency also holds at period $T-t'$. Because

$$\bar{\alpha} \geq \frac{\left[\left(\frac{1}{2}\right)^{T-t'-1} - \left(\frac{1}{2}\right)^{T-t'} \delta\right] p_0}{\left[\left(\frac{1}{2}\right)^{T-t'-1} - \left(\frac{1}{2}\right)^{T-t'} \delta\right] p_0 + (1-\delta)p_1} > \frac{\left[\left(\frac{1}{2}\right)^{T-t'} - \left(\frac{1}{2}\right)^{T-t'+1} \delta\right] p_0}{\left[\left(\frac{1}{2}\right)^{T-t'} - \left(\frac{1}{2}\right)^{T-t'+1} \delta\right] p_0 + (1-\delta)p_1},$$

level-2 players will choose $C$ at period $T-t'+1$ by induction hypothesis. Therefore, it is optimal to choose $C$ at period $T-t'$ if

$$\mathbb{E}u_i^2(C|X,T-t') \geq \left[\frac{\left(\frac{1}{2}\right)^{T-t'} p_0 + dp_1}{\left(\frac{1}{2}\right)^{T-t'-1} p_0 + dp_1}\right]\mathbb{E}u_i^2(C|X,T-t'+1)$$

$$\iff \bar{\alpha} \;\geq\; \frac{\left[\left(\frac{1}{2}\right)^{T-t'-1} - \left(\frac{1}{2}\right)^{T-t'}\delta\right]p_0}{\left[\left(\frac{1}{2}\right)^{T-t'-1} - \left(\frac{1}{2}\right)^{T-t'}\delta\right]p_0 + (1-\delta)p_1}.$$

This completes the proof of sufficiency.

***Step 5:*** Step 4 establishes the base case for $k=2$. Now, suppose there is $K>2$ such that the statement holds for all $2 \leq k \leq K$. It suffices to prove the statement holds for level-$(K+1)$ players. The proof for period $T$ is straightforward. From step 3, we know if $\sigma_i^K(X,T)=1$, then $\sigma_i^{K+1}(X,T)=1$. Hence, the only case that needs to be considered is when

$$\bar{\alpha} \;<\; \frac{\left(\frac{1}{2}\right)^{T-1} p_0}{\left(\frac{1}{2}\right)^{T-1} p_0 + \sum_{j=1}^{K-1} p_j}.$$

By induction hypothesis, $\sigma_{-i}^l(X,t)=0$ for all $1 \leq l \leq K$ and for all $1 \leq t \leq T$. Therefore, $\sigma_i^{K+1}(X,T)=1$ if and only if $\mathbb{E}u_i^{K+1}(C|X,T) \geq 0$, which is equivalent to

$$\bar{\alpha} \;\geq\; \frac{\left(\frac{1}{2}\right)^{T-1} p_0}{\left(\frac{1}{2}\right)^{T-1} p_0 + \sum_{j=1}^{K} p_j}.$$

For any period $2 \leq t \leq T-1$, I first prove the direction of necessity. If

$$\bar{\alpha} \;<\; \frac{\left[\left(\frac{1}{2}\right)^{t-1} - \left(\frac{1}{2}\right)^{t}\delta\right]p_0}{\left[\left(\frac{1}{2}\right)^{t-1} - \left(\frac{1}{2}\right)^{t}\delta\right]p_0 + (1-\delta)\sum_{j=1}^{K} p_j},$$

then by induction hypothesis, $\sigma_{-i}^l(X,t')=0$ for all $1 \leq l \leq K$ and $1 \leq t' \leq t$, implying that $\mathcal{M}_{i+}^{K+1}(t)=\{1,\ldots,K\}$. Then equation (O.2) suggests that the expected payoff of $C$ at period $t$ is

$$\frac{\delta^{t-1}}{\left(\frac{1}{2}\right)^{t-1} p_0 + d\sum_{j=1}^{K} p_j}\left\{d\alpha\left[\left(\frac{1}{2}\right)^{t-1} p_0 + \sum_{j=1}^{K} p_j\right] - (1-d)\left[\left(\frac{1}{2}\right)^{t-1} p_0\right]\right\}.$$

Furthermore, equation (O.3) suggests level-$(K+1)$ players believe the other player will wait at period $t$ with probability

$$\frac{1}{2}\mu^{K+1}(0|X,t) + \sum_{j=1}^{K}\mu^{K+1}(l|X,t) = \frac{\left(\frac{1}{2}\right)^{t} p_0 + d\sum_{j=1}^{K} p_j}{\left(\frac{1}{2}\right)^{t-1} p_0 + d\sum_{j=1}^{K} p_j}.$$

Therefore, by similar calculation as in step 4, choosing $C$ is strictly dominated if

$$\frac{\delta^{t-1}}{\left(\frac{1}{2}\right)^{t-1} p_0 + d\sum_{j=1}^{K} p_j}\left\{d\alpha\left[\left(\frac{1}{2}\right)^{t-1} p_0 + \sum_{j=1}^{K} p_j\right] - (1-d)\left[\left(\frac{1}{2}\right)^{t-1} p_0\right]\right\}$$
$$< \frac{\delta^{t}}{\left(\frac{1}{2}\right)^{t-1} p_0 + d\sum_{j=1}^{K} p_j}\left\{d\alpha\left[\left(\frac{1}{2}\right)^{t} p_0 + \sum_{j=1}^{K} p_j\right] - (1-d)\left[\left(\frac{1}{2}\right)^{t} p_0\right]\right\},$$

which is implied by

$$\bar{\alpha} \;<\; \frac{\left[\left(\frac{1}{2}\right)^{t-1} - \left(\frac{1}{2}\right)^{t}\delta\right] p_0}{\left[\left(\frac{1}{2}\right)^{t-1} - \left(\frac{1}{2}\right)^{t}\delta\right] p_0 + (1-\delta)\sum_{j=1}^{K} p_j}.$$

This proves the direction of necessity.

Second, the sufficiency is proven by induction on the periods. Namely, I show the sufficiency holds for any period $T - t'$ where $1 \leq t' \leq T-2$. Consider the base case for period $T-1$. By step 3, if $\sigma_i^K(X, T-1) = 1$, then $\sigma_i^{K+1}(X, T-1) = 1$. Therefore, it suffices to consider the case where

$$\frac{\left[\left(\frac{1}{2}\right)^{T-2} - \left(\frac{1}{2}\right)^{T-1}\delta\right] p_0}{\left[\left(\frac{1}{2}\right)^{T-2} - \left(\frac{1}{2}\right)^{T-1}\delta\right] p_0 + (1-\delta)\sum_{j=1}^{K} p_j} \;\leq\; \bar{\alpha} \;<\; \frac{\left[\left(\frac{1}{2}\right)^{T-2} - \left(\frac{1}{2}\right)^{T-1}\delta\right] p_0}{\left[\left(\frac{1}{2}\right)^{T-2} - \left(\frac{1}{2}\right)^{T-1}\delta\right] p_0 + (1-\delta)\sum_{j=1}^{K-1} p_j}.$$

By induction hypothesis, $\sigma_{-i}^l(X, t) = 0$ for all $1 \leq t \leq T-1$ and $1 \leq l \leq K$. Moreover, $\sigma_i^{K+1}(X, T) = 1$ because

$$\frac{\left[\left(\frac{1}{2}\right)^{T-2} - \left(\frac{1}{2}\right)^{T-1}\delta\right] p_0}{\left[\left(\frac{1}{2}\right)^{T-2} - \left(\frac{1}{2}\right)^{T-1}\delta\right] p_0 + (1-\delta)\sum_{j=1}^{K} p_j} > \frac{\left(\frac{1}{2}\right)^{T-1} p_0}{\left(\frac{1}{2}\right)^{T-1} p_0 + \sum_{j=1}^{K} p_j}.$$

Therefore, by a similar calculation as in step 4, it can be proven that it is optimal for level-$(K+1)$ players to choose $C$ at period $T-1$ if

$$\mathbb{E}u_i^{K+1}(C|X, T-1) \geq \left[\frac{\left(\frac{1}{2}\right)^{T-1} p_0 + d\sum_{j=1}^{K} p_j}{\left(\frac{1}{2}\right)^{T-2} p_0 + d\sum_{j=1}^{K} p_j}\right]\mathbb{E}u_i^{K+1}(C|X, T)$$
$$\iff \bar{\alpha} \;\geq\; \frac{\left[\left(\frac{1}{2}\right)^{T-2} - \left(\frac{1}{2}\right)^{T-1}\delta\right] p_0}{\left[\left(\frac{1}{2}\right)^{T-2} - \left(\frac{1}{2}\right)^{T-1}\delta\right] p_0 + (1-\delta)\sum_{j=1}^{K} p_j}.$$

Now, suppose there is $t' \leq T-2$ such that the statement holds for any period $T-t$ where $1 \leq t \leq t'-1$. It can be proven that the statement also holds at period period $T-t'$. By step 3, if $\sigma_i^K(X, T-t') = 1$, then $\sigma_i^{K+1}(X, T-t') = 1$ and it suffices to consider the case:

$$\frac{\left[\left(\frac{1}{2}\right)^{T-t'-1} - \left(\frac{1}{2}\right)^{T-t'}\delta\right] p_0}{\left[\left(\frac{1}{2}\right)^{T-t'-1} - \left(\frac{1}{2}\right)^{T-t'}\delta\right] p_0 + (1-\delta)\sum_{j=1}^{K} p_j} \;\leq\; \bar{\alpha} < \frac{\left[\left(\frac{1}{2}\right)^{T-t'-1} - \left(\frac{1}{2}\right)^{T-t'}\delta\right] p_0}{\left[\left(\frac{1}{2}\right)^{T-t'-1} - \left(\frac{1}{2}\right)^{T-t'}\delta\right] p_0 + (1-\delta)\sum_{j=1}^{K-1} p_j}.$$

By induction hypothesis, $\sigma_{-i}^{l}(X,t)=0$ for all $1\leq t\leq T-t'$ and $1\leq l\leq K$, and $\sigma_i^{K+1}(X,T-t'+1)=1$. Therefore, by a similar calculation as in step 4, it can be proven that it is optimal for level-$(K+1)$ players to choose $C$ at period $T-t'$ if

$$\bar{\alpha} \;\geq\; \frac{\left[\left(\frac{1}{2}\right)^{T-t'-1}-\left(\frac{1}{2}\right)^{T-t'}\delta\right]p_0}{\left[\left(\frac{1}{2}\right)^{T-t'-1}-\left(\frac{1}{2}\right)^{T-t'}\delta\right]p_0+(1-\delta)\sum_{j=1}^{K}p_j}.$$

This completes the proof of the proposition. ■

## Proof of Corollary 1

By Proposition 3, we know $\sigma_i^k(O,t)=1$ for all $t$ and $k\geq 1$, and $\sigma_i^1(X,t)=0$ for all $t$. Then by Definition 1, we can obtain that $\hat{\sigma}_i^k(O)=1$ for all $k\geq 1$, and $\hat{\sigma}_i^1(X)=T+1$. In addition, since $\sigma_i^k(X,1)=0$ for all $k\geq 2$, $\hat{\sigma}_i^k(X)\neq 1$. Moreover, the DCH solution can be equivalently characterized by optimal stopping periods because for any $t\geq 2$ and $k\geq 2$,

$$\begin{aligned}
\hat{\sigma}_i^k(X)=t &\iff \sigma_i^k(X,t-1)=0 \quad\text{and}\quad \sigma_i^k(X,t)=1,\\
\hat{\sigma}_i^k(X)=T+1 &\iff \sigma_i^k(X,t')=0 \quad\text{for any } 1\leq t'\leq T.
\end{aligned}$$

Lastly, to show the monotonicity, it suffices to show that for any $k'>k\geq 2$ and any $2\leq t\leq T$, if $\sigma_i^k(X,t)=1$, then $\sigma_i^{k'}(X,t)=1$. The discussion is separated into two cases. First, if $t=T$, then by Proposition 3, $\sigma_i^k(X,T)=1$ suggests

$$\bar{\alpha} \;\geq\; \frac{\left(\frac{1}{2}\right)^{T-1}p_0}{\left(\frac{1}{2}\right)^{T-1}p_0+\sum_{j=1}^{k-1}p_j} \;>\; \frac{\left(\frac{1}{2}\right)^{T-1}p_0}{\left(\frac{1}{2}\right)^{T-1}p_0+\sum_{j=1}^{k'-1}p_j},$$

implying $\sigma_i^{k'}(X,T)=1$. Second, if $2\leq t\leq T-1$, by Proposition 3, $\sigma_i^k(X,t)=1$ suggests

$$\bar{\alpha} \;\geq\; \frac{\left[\left(\frac{1}{2}\right)^{t-1}-\left(\frac{1}{2}\right)^{t}\delta\right]p_0}{\left[\left(\frac{1}{2}\right)^{t-1}-\left(\frac{1}{2}\right)^{t}\delta\right]p_0+(1-\delta)\sum_{j=1}^{k-1}p_j} \;>\; \frac{\left[\left(\frac{1}{2}\right)^{t-1}-\left(\frac{1}{2}\right)^{t}\delta\right]p_0}{\left[\left(\frac{1}{2}\right)^{t-1}-\left(\frac{1}{2}\right)^{t}\delta\right]p_0+(1-\delta)\sum_{j=1}^{k'-1}p_j},$$

implying $\sigma_i^{k'}(X,t)=1$. This completes the proof. ■

## Proof of Proposition 4

***Step 1:*** Consider any $i\in N$. If $x_{-i}=O$, player $i$ knows his face is dirty immediately, suggesting 1 is a dominant strategy and $\tilde{\sigma}_i^k(O)=1$ for any $k\geq 1$. If $x_{-i}=X$, the expected payoff of 1 is $d\alpha-(1-d)<0$, implying $\tilde{\sigma}_i^k(X)\geq 2$ for any $k\geq 1$. Moreover, level-1 players believe the other player is level-0, so when observing $X$, the expected payoff of $2\leq j\leq T$ is

$$d\left[\frac{T+2-j}{T+1}\delta^{j-1}\alpha\right]-(1-d)\left[\frac{T+2-j}{T+1}\delta^{j-1}\right]=\delta^{j-1}\left(\frac{T+2-j}{T+1}\right)[d\alpha-(1-d)]<0.$$

implying $\tilde{\sigma}_i^1(X) = T+1$.

***Step 2:*** This step proves for any $K > 1$, if $\tilde{\sigma}_i^{l+1}(X) \leq \tilde{\sigma}_i^l(X)$ for all $1 \leq l \leq K-1$, then $\tilde{\sigma}_i^{K+1}(X) \leq \tilde{\sigma}_i^K(X)$. Note that if $\tilde{\sigma}_i^K(X) = T+1$, then there is nothing to prove. Let $s^* \equiv \tilde{\sigma}_i^K(X)$ and focus on the case where $2 \leq s^* \leq T$. If $s^* = T$, then level-$(K+1)$ players' expected payoff of choosing $T$ is

$$\begin{aligned}
&\delta^{T-1}\left[d\alpha\left(\frac{2}{T+1}\frac{p_0}{\sum_{j=0}^{K}p_j} + \frac{\sum_{j=1}^{K}p_j}{\sum_{j=0}^{K}p_j}\right) - (1-d)\left(\frac{2}{T+1}\frac{p_0}{\sum_{j=0}^{K}p_j}\right)\right] \\
&> \delta^{T-1}\left[d\alpha\left(\frac{2}{T+1}\frac{p_0}{\sum_{j=0}^{K-1}p_j} + \frac{\sum_{j=1}^{K-1}p_j}{\sum_{j=0}^{K-1}p_j}\right) - (1-d)\left(\frac{2}{T+1}\frac{p_0}{\sum_{j=0}^{K-1}p_j}\right)\right] \geq 0.
\end{aligned}$$

The last inequality holds because it is optimal for level-$K$ players to choose $T$. This suggests that $T+1$ is dominated by $T$ and hence $\tilde{\sigma}_i^{K+1}(X) \leq T = \tilde{\sigma}_i^K(X)$.

Next, consider $2 \leq s^* \leq T-1$. If level-$(K+1)$ player $i$ chooses some $s$ where $s^* < s < T+1$ that yields a non-negative expected payoff, then choosing $s$ is strictly suboptimal because

$$\begin{aligned}
&\delta^{s-1}\left[d\alpha\left(\frac{T+2-s}{T+1}\frac{p_0}{\sum_{j=0}^{K}p_j} + \frac{\sum_{j=1}^{K-1}p_j}{\sum_{j=0}^{K}p_j}\right) - (1-d)\left(\frac{T+2-s}{T+1}\frac{p_0}{\sum_{j=0}^{K}p_j}\right)\right] \\
&< \delta^{s-1}\left[d\alpha\left(\frac{T+2-s}{T+1}\frac{p_0}{\sum_{j=0}^{K-1}p_j} + \frac{\sum_{j=1}^{K-1}p_j}{\sum_{j=0}^{K-1}p_j}\right) - (1-d)\left(\frac{T+2-s}{T+1}\frac{p_0}{\sum_{j=0}^{K-1}p_j}\right)\right] \\
&\leq \delta^{s^*-1}\left[d\alpha\left(\frac{T+2-s^*}{T+1}\frac{p_0}{\sum_{j=0}^{K-1}p_j} + \frac{\sum_{j=1}^{K-1}p_j}{\sum_{j=0}^{K-1}p_j}\right) - (1-d)\left(\frac{T+2-s^*}{T+1}\frac{p_0}{\sum_{j=0}^{K-1}p_j}\right)\right] \\
&< \delta^{s^*-1}\left[d\alpha\left(\frac{T+2-s^*}{T+1}\frac{p_0}{\sum_{j=0}^{K}p_j} + \frac{\sum_{j=1}^{K}p_j}{\sum_{j=0}^{K}p_j}\right) - (1-d)\left(\frac{T+2-s^*}{T+1}\frac{p_0}{\sum_{j=0}^{K}p_j}\right)\right].
\end{aligned}$$

Note that the second inequality holds because $s^*$ is level-$K$ players' optimal choice, and the RHS of the last inequality is level-$(K+1)$ players' expected payoff of choosing $s^*$. These inequalities show that it is not optimal for level-$(K+1)$ players to choose any $s > s^*$, suggesting that $\tilde{\sigma}_i^{K+1}(X) \leq \tilde{\sigma}_i^K(X)$.

***Step 3:*** The proposition is proven by induction on $k$. This step establishes the base case for level-2 players. For any $2 \leq j \leq T$, the expected payoff of choosing $j$ is $\mathbb{E}u_i^2(j|X) =$

$$d\left[\left(\frac{T+2-j}{T+1}\delta^{j-1}\alpha\right)\frac{p_0}{p_0+p_1} + \left(\delta^{j-1}\alpha\right)\frac{p_1}{p_0+p_1}\right] - (1-d)\left[\left(\frac{T+2-j}{T+1}\delta^{j-1}\right)\frac{p_0}{p_0+p_1}\right].$$

For level-2 players and any $2 \leq j \leq T-1$, let $\Delta_j^2 \equiv \mathbb{E}u_i^2(j|X) - \mathbb{E}u_i^2(j+1|X)$ be the difference of expected payoffs between $j$ and $j+1$. That is,

$$\begin{aligned}
\Delta_j^2 = \delta^{j-1}d\alpha&\left[\left(\frac{T+2-j}{T+1} - \frac{T+1-j}{T+1}\delta\right)\frac{p_0}{p_0+p_1} + (1-\delta)\frac{p_1}{p_0+p_1}\right] \\
&- \delta^{j-1}(1-d)\left[\left(\frac{T+2-j}{T+1} - \frac{T+1-j}{T+1}\delta\right)\frac{p_0}{p_0+p_1}\right],
\end{aligned}$$

suggesting that $j$ dominates $j+1$ if and only if

$$\Delta_j^2 \geq 0 \iff \bar{\alpha} \geq \frac{\left[\frac{T+2-j}{T+1} - \frac{T+1-j}{T+1}\delta\right] p_0}{\left[\frac{T+2-j}{T+1} - \frac{T+1-j}{T+1}\delta\right] p_0 + (1-\delta)p_1}.$$

Because the RHS is decreasing function in $j$, $\Delta_j^2 \geq 0$ implies $\Delta_{j+1}^2 \geq 0$. Moreover, since

$$\mathbb{E}u_i^2(j|X) \geq 0 \iff \bar{\alpha} \ \geq \ \frac{\frac{T+2-j}{T+1} p_0}{\frac{T+2-j}{T+1} p_0 + p_1},$$

$\Delta_j^2 \geq 0$ implies $\mathbb{E}u_i^2(j|X) \geq 0$ because

$$\bar{\alpha} \ \geq \ \frac{\left[\frac{T+2-j}{T+1} - \frac{T+1-j}{T+1}\delta\right] p_0}{\left[\frac{T+2-j}{T+1} - \frac{T+1-j}{T+1}\delta\right] p_0 + (1-\delta)p_1} \ > \ \frac{\frac{T+2-j}{T+1} p_0}{\frac{T+2-j}{T+1} p_0 + p_1}.$$

As a result, $\tilde{\sigma}_i^2(X) \leq T$ if and only if $\mathbb{E}u_i^2(T|X) \geq 0$, which is equivalent to

$$\bar{\alpha} \ \geq \ \frac{\frac{2}{T+1} p_0}{\frac{2}{T+1} p_0 + p_1},$$

and for any other $2 \leq t \leq T-1$, $\tilde{\sigma}_i^2(X) \leq t$ if and only if

$$\Delta_t^2 \geq 0 \iff \bar{\alpha} \ \geq \ \frac{\left[\frac{T+2-t}{T+1} - \frac{T+1-t}{T+1}\delta\right] p_0}{\left[\frac{T+2-t}{T+1} - \frac{T+1-t}{T+1}\delta\right] p_0 + (1-\delta)p_1}.$$

***Step 4:*** Step 3 establishes the base case where $k = 2$. Now suppose there is $K > 2$ such that the statement holds for any $2 \leq k \leq K$. It suffices to prove that the statement also holds for level-$(K+1)$ players. By step 1, $\tilde{\sigma}_i^{K+1}(X) \geq 2$. Besides, note that for any $1 \leq t \leq T$ and $1 \leq l \leq K$, if $\tilde{\sigma}_{-i}^l(X) > t$, then level-$(K+1)$ player $i$'s expected payoff of choosing $2 \leq j \leq t+1$ is $\mathbb{E}u_i^{K+1}(j|X) =$

$$\delta^{j-1}\left[d\alpha\left(\frac{T+2-j}{T+1}\frac{p_0}{\sum_{j=0}^K p_j} + \frac{\sum_{j=1}^K p_j}{\sum_{j=0}^K p_j}\right) - (1-d)\left(\frac{T+2-j}{T+1}\frac{p_0}{\sum_{j=0}^K p_j}\right)\right].$$

Similar to step 3, we define $\Delta_{t'}^{K+1}$ for any $2 \leq t' \leq t$ where $\Delta_{t'}^{K+1}$ is the difference of expected payoff between choosing $t'$ and $t'+1$. That is,

$$\begin{aligned}\Delta_{t'}^{K+1} \equiv \delta^{t'-1} d\alpha &\left[\left(\frac{T+2-t'}{T+1} - \frac{T+1-t'}{T+1}\delta\right)\frac{p_0}{\sum_{j=0}^K p_j} + (1-\delta)\frac{\sum_{j=1}^K p_j}{\sum_{j=0}^K p_j}\right] \\ &- \delta^{t'-1}(1-d)\left[\left(\frac{T+2-t'}{T+1} - \frac{T+1-t'}{T+1}\delta\right)\frac{p_0}{\sum_{j=0}^K p_j}\right].\end{aligned}$$

By the same argument as in step 3, $\Delta_{t'}^{K+1} < 0$ implies $\Delta_{t'-1}^{K+1} < 0$. Therefore, if $\tilde{\sigma}_{-i}^{l}(X) > t$ for any $1 \leq l \leq K$, it is strictly dominated for level-$(K+1)$ players to choose $t'$ (and all strategies $s < t'$) where $2 \leq t' \leq t$ if

$$\bar{\alpha} \; < \; \frac{\left[\frac{T+2-t'}{T+1} - \frac{T+1-t'}{T+1}\delta\right] p_0}{\left[\frac{T+2-t'}{T+1} - \frac{T+1-t'}{T+1}\delta\right] p_0 + (1-\delta)\sum_{j=1}^{K} p_j}, \tag{O.4}$$

and by a similar argument as in step 3, $\Delta_{t'}^{K+1} \geq 0$ implies $\mathbb{E}u_i^{K+1}(t'|X) \geq 0$.

The proof for period $T$ is straightforward. The implication of the induction hypothesis is that $\tilde{\sigma}_i^{l+1}(X) \leq \tilde{\sigma}_i^{l}(X)$ for all $1 \leq l \leq K-1$. By step 2, $\tilde{\sigma}_i^{K+1}(X) \leq T$ if $\tilde{\sigma}_i^{K}(X) \leq T$. Thus, it suffices to consider the case where

$$\bar{\alpha} \; < \; \frac{\frac{2}{T+1}p_0}{\frac{2}{T+1}p_0 + \sum_{j=1}^{K-1} p_j}.$$

By induction hypothesis, $\tilde{\sigma}_i^{l}(X) = T+1$ for all $1 \leq l \leq K$, so $\tilde{\sigma}_i^{K+1}(X) \leq T$ if and only if

$$\mathbb{E}u_i^{K+1}(T|X) \geq 0 \iff \bar{\alpha} \; \geq \; \frac{\frac{2}{T+1}p_0}{\frac{2}{T+1}p_0 + \sum_{j=1}^{K} p_j}.$$

Next, consider any $2 \leq t \leq T-1$. By induction hypothesis and step 2, if $\tilde{\sigma}_i^{K}(X) \leq t$, then $\tilde{\sigma}_i^{K+1}(X) \leq t$. Hence, it suffices to complete the proof by considering

$$\bar{\alpha} \; < \; \frac{\left[\frac{T+2-t}{T+1} - \frac{T+1-t}{T+1}\delta\right] p_0}{\left[\frac{T+2-t}{T+1} - \frac{T+1-t}{T+1}\delta\right] p_0 + (1-\delta)\sum_{j=1}^{K-1} p_j}.$$

In this case, $t < \tilde{\sigma}_i^{l+1}(X) \leq \tilde{\sigma}_i^{l}(X)$ for all $1 \leq l \leq K-1$. Therefore, inequality (O.4) implies that $\tilde{\sigma}_i^{K+1}(X) \leq t$ if and only if

$$\bar{\alpha} \; \geq \; \frac{\left[\frac{T+2-t}{T+1} - \frac{T+1-t}{T+1}\delta\right] p_0}{\left[\frac{T+2-t}{T+1} - \frac{T+1-t}{T+1}\delta\right] p_0 + (1-\delta)\sum_{j=1}^{K} p_j}.$$

This completes the proof of this proposition. ■

## Proof of Corollary 2

It suffices to prove the monotonicity by showing for all $k' > k \geq 2$, if $\tilde{\sigma}_i^{k}(X) \leq t$, then $\tilde{\sigma}_i^{k'}(X) \leq t$ for any $2 \leq t \leq T$. We can separate the analysis into two cases. First, if $t = T$, then by Proposition 4, $\tilde{\sigma}_i^{k}(X) \leq T$ suggests

$$\bar{\alpha} \; \geq \; \frac{\frac{2}{T+1}p_0}{\frac{2}{T+1}p_0 + \sum_{j=1}^{k-1} p_j} \; > \; \frac{\frac{2}{T+1}p_0}{\frac{2}{T+1}p_0 + \sum_{j=1}^{k'-1} p_j},$$

implying $\tilde{\sigma}_i^{k'}(X) \leq T$. Second, for any $2 \leq t \leq T-1$, by Proposition 4, $\tilde{\sigma}_i^{k}(X) \leq t$ suggests

$$\bar{\alpha} \; \geq \; \frac{\left[\frac{T+2-t}{T+1} - \frac{T+1-t}{T+1}\delta\right] p_0}{\left[\frac{T+2-t}{T+1} - \frac{T+1-t}{T+1}\delta\right] p_0 + (1-\delta)\sum_{j=1}^{k-1} p_j} \; > \; \frac{\left[\frac{T+2-t}{T+1} - \frac{T+1-t}{T+1}\delta\right] p_0}{\left[\frac{T+2-t}{T+1} - \frac{T+1-t}{T+1}\delta\right] p_0 + (1-\delta)\sum_{j=1}^{k'-1} p_j},$$

implying $\tilde{\sigma}_i^{k'}(X) \leq t$. This completes the proof. ■

## Proof of Proposition 5

First, for any $k \geq 2$, it suffices to prove $\mathcal{S}_T^k \subset \mathcal{E}_T^k$ by showing if $\tilde{\sigma}_i^k(X) \leq T$, then $\hat{\sigma}_i^k(X) \leq T$. This is true because

$$\frac{\frac{2}{T+1}p_0}{\frac{2}{T+1}p_0 + \sum_{j=1}^{k-1} p_j} > \frac{\left(\frac{1}{2}\right)^{T-1} p_0}{\left(\frac{1}{2}\right)^{T-1} p_0 + \sum_{j=1}^{k-1} p_j}.$$

Similarly, for $2 \leq t \leq T-1$, we can first observe that $\mathcal{S}_t^k \subset \mathcal{E}_t^k$ if and only if

$$\frac{\left[\frac{T+2-t}{T+1} - \frac{T+1-t}{T+1}\delta\right] p_0}{\left[\frac{T+2-t}{T+1} - \frac{T+1-t}{T+1}\delta\right] p_0 + (1-\delta)\sum_{j=1}^{k-1} p_j} \geq \frac{\left[\left(\frac{1}{2}\right)^{t-1} - \left(\frac{1}{2}\right)^{t}\delta\right] p_0}{\left[\left(\frac{1}{2}\right)^{t-1} - \left(\frac{1}{2}\right)^{t}\delta\right] p_0 + (1-\delta)\sum_{j=1}^{k-1} p_j}$$

$$\iff \delta \leq \frac{(2^t-2)(T+1)-(t-1)2^t}{(2^t-1)(T+1)-t2^t} \equiv \overline{\delta}(T,t). \tag{O.5}$$

where $\overline{\delta}(T,t) > 0$ because $(2^t-2)(T+1)-(t-1)2^t \geq 2(T+1)-4 > 0$ and $(2^t-1)(T+1)-t2^t \geq 3(T+1)-8 > 0$. If $\overline{\delta}(T,t) > 1$, then the inequality holds for any $\delta \in (0,1)$, and hence $\mathcal{S}_t^k \subset \mathcal{E}_t^k$. Otherwise, if $\overline{\delta}(T,t) < 1$, the inequality does not hold for all $\delta$, implying there is no set inclusion relationship between $\mathcal{S}_t^k$ and $\mathcal{E}_t^k$. In addition, inequality (O.5) suggests $\hat{\sigma}_i^k(X) < \tilde{\sigma}_i^k(X)$ if $\delta < \overline{\delta}(T,t)$ and $\hat{\sigma}_i^k(X) > \tilde{\sigma}_i^k(X)$ if $\delta > \overline{\delta}(T,t)$. Lastly, as we rearrange the inequality, we can obtain that

$$\overline{\delta}(T,t) < 1 \iff \frac{(2^t-2)(T+1)-(t-1)2^t}{(2^t-1)(T+1)-t2^t} < 1 \iff t < \frac{\ln(T+1)}{\ln(2)}.$$

This completes the proof of this proposition. ■

## Proof of Corollary 3

By Proposition 5, we know for any $k \geq 2$, there is no set inclusion relationship between $\mathcal{S}_t^k$ and $\mathcal{E}_t^k$ if $2 \leq t < [\ln(T+1)/\ln(2)]$. When $T \to \infty$, this condition holds for any $t \geq 2$. Moreover, from Proposition 5, we can obtain that

$$\overline{\delta}^*(t) = \lim_{T\to\infty} \overline{\delta}(T,t) = \lim_{T\to\infty} \frac{(2^t-2)(T+1)-(t-1)2^t}{(2^t-1)(T+1)-t2^t} = \frac{2^t-2}{2^t-1}.$$

This completes the proof. ■

# 2 Additional Analysis

## 2.1 Additional Result for Poisson-DCH

One feature of the Poisson-DCH model is that as $\tau \to \infty$, the aggregate choice frequencies converge to the equilibrium prediction. This provides a second interpretation for the parameter $\tau$: the higher the value of $\tau$, the closer the predictions are to the equilibrium. It is worth noting that, as highlighted by Camerer et al. (2004), this convergence property does not hold for general classes of games.

For the sake of simplicity, I will prove the result for sequential two-person games. A similar argument holds for the simultaneous version. For any two-person dirty faces game, conditional on there is an announcement, there are two possible states: one dirty face or two dirty faces, which are denoted as $\Omega = \{OX, XX\}$. For each $\omega \in \Omega$, equilibrium predicts a deterministic terminal period. We use $F^*_\omega(t)$ to express the (degenerated) distribution of terminal periods at the equilibrium. The equilibrium predicts that players will choose $C$ at period 1 when seeing $O$, and choose $W$ at period 1 and $C$ at period 2 when seeing $X$. Therefore, when $\omega = OX$, the game will end at period 1, and when $\omega = XX$, the game will end at period 2. In other words,

$$F^*_{OX}(t) = \begin{cases} 0 & \text{if } t < 1 \\ 1 & \text{if } t \geq 1, \end{cases} \qquad \text{and} \qquad F^*_{XX}(t) = \begin{cases} 0 & \text{if } t < 2 \\ 1 & \text{if } t \geq 2. \end{cases}$$

In contrast, given any $\tau > 0$ and $\omega \in \Omega$, the Poisson-DCH model predicts a non-degenerated distribution over all possible terminal periods. We use $F^D_w(t|\tau)$ to denote the distribution predicted by the Poisson-DCH. Proposition O.1 states that when $\tau \to \infty$, the max norm between $F^D_\omega(t|\tau)$ and $F^*_\omega(t)$ will converge to 0 for any $\omega \in \Omega$.

**Proposition O.1.** *Consider any sequential two-person dirty faces game. When the prior distribution of levels follows Poisson($\tau$), for any $\omega \in \Omega$, $\lim_{\tau\to\infty} \left\|F^*_\omega(t) - F^D_\omega(t|\tau)\right\|_\infty = 0$.*

*Proof.* When $\omega = OX$, the strategic player that sees a clean face will choose $C$ in period 1. Thus,

$$F^D_{OX}(1|\tau) = 1 - \left(\frac{1}{2}e^{-\tau}\right)\left(1 - \frac{1}{2}e^{-\tau}\right).$$

To show $\left\|F^*_{OX}(t) - F^D_{OX}(t|\tau)\right\|_\infty \to 0$, it suffices to show $F^D_{OX}(1|\tau) \to 1$ as $\tau \to \infty$, which holds since

$$\lim_{\tau\to\infty} F^D_{OX}(1|\tau) = \lim_{\tau\to\infty} 1 - \left(\frac{1}{2}e^{-\tau}\right)\left(1 - \frac{1}{2}e^{-\tau}\right) = 1.$$

When $\omega = XX$, it suffices to prove the convergence by showing $F^D_{XX}(1|\tau) \to 0$ and $F^D_{XX}(2|\tau) \to 1$ as $\tau \to \infty$. Since every level $k \geq 1$ will choose $W$ in period 1 when seeing a dirty face, $F^D_{XX}(1|\tau) = 1 - [1 - (1/2)e^{-\tau}]^2$, implying that

$$\lim_{\tau\to\infty} F^D_{XX}(1|\tau) = \lim_{\tau\to\infty} 1 - \left[1 - \frac{1}{2}e^{-\tau}\right]^2 = 0.$$

Lastly, let $K^*(\tau)$ be the lowest level of players to choose $C$ at period 2 when seeing a dirty face with the prior distribution of levels being Poisson($\tau$). By Proposition 3, $K^*(\tau)$ is weakly decreasing in $\tau$, and $K^*(\tau) \to 2$ as $\tau \to \infty$. Hence,

$$F^D_{XX}(2|\tau) = 1 - \left[(1/4)e^{-\tau} + \sum_{j=1}^{K^*(\tau)-1} e^{-\tau}\tau^j/j!\right]^2,$$

suggesting the limit is

$$\lim_{\tau\to\infty} F^D_{XX}(2|\tau) = \lim_{\tau\to\infty} 1 - \left[\frac{1}{4}e^{-\tau} + \sum_{j=1}^{K^*(\tau)-1} \frac{e^{-\tau}\tau^j}{j!}\right]^2 = \lim_{\tau\to\infty} 1 - \left[\frac{1}{4}e^{-\tau} + \tau e^{-\tau}\right]^2 = 1. \blacksquare$$

## 2.2 Supplementary Tables

This section includes all the supplementary tables from the experiment. Table O.1 lists the empirical frequencies of choosing $C$ at each information set for both treatments. Table O.2 presents the empirical frequencies of choosing $C$ at information set $(2, X)$ for different payoff structures and treatments.

Table O.1: The Empirical Frequencies of $C$ at Each Information Set

| $x_{-i}$ | Sequential Treatment | | | | | | Simultaneous Treatment | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $O$ | | | $X$ | | | $O$ | | | $X$ | | |
| | Obs | Claim % | s.d. | Obs | Claim % | s.d. | Obs | Claim % | s.d. | Obs | Claim % | s.d. |
| Periods | | | | | | | | | | | | |
| 1 | 148 | 0.845 | 0.364 | 572 | 0.313 | 0.464 | 148 | 0.811 | 0.393 | 548 | 0.263 | 0.441 |
| 2 | 16 | 0.438 | 0.512 | 210 | 0.600 | 0.491 | 28 | 0.250 | 0.441 | 404 | 0.223 | 0.417 |
| 3 | 4 | 0.000 | 0.000 | 34 | 0.206 | 0.410 | 21 | 0.190 | 0.402 | 314 | 0.172 | 0.378 |
| 4 | 3 | 0.000 | 0.000 | 21 | 0.190 | 0.402 | 16 | 0.250 | 0.447 | 259 | 0.131 | 0.338 |
| 5 | 2 | 0.500 | 0.707 | 14 | 0.214 | 0.426 | 14 | 0.143 | 0.363 | 227 | 0.172 | 0.378 |

Note: For the simultaneous treatment, the choice data at the information set level are implied by the contingent strategies. For instance, choosing the contingent strategy "claim at period 4" implies that the subject will wait from period 1 to period 3 and claim in period 4.

Table O.2: The Empirical Frequencies of $C$ at Information Set $(X, 2)$ for Different Games

| | Sequential Treatment | | | Simultaneous Treatment | | |
|---|---|---|---|---|---|---|
| $(\delta, \bar{\alpha})$ | Obs | Claim % | s.d. | Obs | Claim % | s.d. |
| Diagnostic Games | | | | | | |
| (0.60, 0.45) | 39 | 0.564 | 0.502 | 78 | 0.256 | 0.439 |
| (0.95, 0.80) | 36 | 0.667 | 0.479 | 59 | 0.237 | 0.429 |
| Control Games | | | | | | |
| (0.60, 0.80) | 38 | 0.789 | 0.413 | 61 | 0.361 | 0.484 |
| (0.80, 0.45) | 35 | 0.543 | 0.505 | 67 | 0.134 | 0.344 |
| (0.80, 0.80) | 24 | 0.542 | 0.509 | 63 | 0.190 | 0.396 |
| (0.95, 0.45) | 38 | 0.474 | 0.506 | 76 | 0.171 | 0.379 |

To compute the measure of violation of invariance under strategic equivalence of each payoff structure $(\delta, \bar{\alpha})$, I run the following regression on the data of information set $(X, 2)$:

$$\mathbb{1}\{\text{claim}\}_i = \alpha_0 + \alpha_1 \mathbb{1}\{\text{sequential}\}_i + \epsilon_i \tag{O.6}$$

where $\mathbb{1}\{\text{claim}\}_i$ is the dummy variable for player $i$ choosing $C$ and $\mathbb{1}\{\text{sequential}\}_i$ is the dummy variable for the sequential treatment. Table O.6 reports the results for all payoff structures. The standard errors are clustered at the session level.

Table O.6: The Magnitude of the Treatment Effect for Different Games

| Payoff Structure $(\delta, \bar{\alpha})$ | (0.60, 0.45) | (0.60, 0.80) | (0.80, 0.45) | (0.80, 0.80) | (0.95, 0.45) | (0.95, 0.80) |
|---|---|---|---|---|---|---|
| Sequential Treatment | 0.308*** | 0.429* | 0.409* | 0.351** | 0.303** | 0.429** |
| | (0.066) | (0.160) | (0.160) | (0.101) | (0.073) | (0.110) |
| Constant | 0.256*** | 0.361*** | 0.134*** | 0.190*** | 0.171** | 0.237** |
| | (0.057) | (0.070) | (0.018) | (0.025) | (0.049) | (0.062) |
| N | 117 | 99 | 102 | 87 | 114 | 95 |
| R-squared | 0.0914 | 0.1744 | 0.1889 | 0.1203 | 0.1028 | 0.1808 |

Note: The standard errors are clustered at the session level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Finally, Table O.7 and O.8 report the distributions of reaction times when players see a dirty face in the sequential and simultaneous treatments, respectively.

Table O.7: Reaction Times (seconds) when Seeing $X$ in the Sequential Treatment

| Periods | Obs | Mean | s.d. | Q1 | Median | Q3 |
|---|---|---|---|---|---|---|
| 1 | 572 | 11.29 | 9.845 | 5.359 | 7.574 | 13.72 |
| 2 | 210 | 8.172 | 6.531 | 4.495 | 6.106 | 9.745 |
| 3 | 34 | 7.530 | 3.816 | 4.785 | 6.277 | 11.44 |
| 4 | 21 | 6.901 | 4.135 | 3.150 | 6.299 | 9.767 |
| 5 | 14 | 7.663 | 5.597 | 3.231 | 6.677 | 8.856 |
| All | 851 | 10.20 | 8.917 | 5.000 | 7.152 | 12.08 |

Table O.8: Reaction Times (seconds) when Seeing $X$ in the Simultaneous Treatment

| Stopping Strategies | Obs | Mean | s.d. | Q1 | Median | Q3 |
|---|---|---|---|---|---|---|
| Claim at 1 | 144 | 11.93 | 8.804 | 5.811 | 8.958 | 15.91 |
| Claim at 2 | 90 | 13.85 | 10.18 | 7.191 | 11.32 | 16.62 |
| Claim at 3 | 54 | 17.64 | 12.81 | 8.095 | 14.30 | 24.61 |
| Claim at 4 | 34 | 21.15 | 15.16 | 12.32 | 16.46 | 22.96 |
| Claim at 5 | 39 | 23.34 | 14.43 | 13.84 | 19.76 | 27.83 |
| Always Wait | 187 | 14.24 | 11.09 | 7.342 | 10.48 | 20.24 |
| All | 548 | 15.04 | 11.58 | 7.413 | 11.32 | 19.48 |

# 3 Experimental Instructions

## 3.1 Sequential Treatment

**General Instructions**

Thank you for participating in the experiment. You are about to take part in a decision-making experiment, in which your earnings will depend partly on your decisions, partly on the decision of others, and partly on chance.

The entire session will take place through computer terminals, and all interactions between participants will be conducted through the computers. Please do not talk or in any way try to communicate with other participants during the session.

The main task of the experiment consists of 12 matches. Before the main task, you will be asked to complete some comprehension questions. If you have any questions, please raise your hand and the question will be answered so that everyone can hear.

In this experiment, you will earn “points” in each match. Your earnings will be determined by the total points you earn in the 12 matches. Each point has a value of $0.02. That is, every 100 points generates $2 in earnings for you. In addition to your earnings from decisions, you will receive a show-up fee of $10. At the end of the experiment, your earnings will be rounded up to the nearest dollar amount. All your earnings will be paid in cash privately at the end of the experiment.

**Main Task**

1. In this experiment, you will be asked to make decisions in 12 matches. You will be randomly matched with another participant into a group for every separate match. This random pairing changes in every match.

2. Each match in this experiment corresponds to a game with the following rules.

- At the beginning of each match, each of you and the other participant will be randomly assigned a "color" (either Red or White). After the colors are assigned, you will be able to see the color of the other participant who is paired with you. However, you cannot see your own color!

- There are 3 possible situations, and the probabilities of these situations are summarized in the following table.

| Situations | Probabilities |
|---|---|
| You are Red and the other participant is White. | $p$ |
| You are White and the other participant is Red. | $p$ |
| Both of you are Red. | $1-2p$ |

  In other words, there is always at least one Red participant among each group.

- Each match is played in rounds. In each match, there are at most 5 rounds. Your color and the other participant's color are fixed in the match. In each round, you and the other participant will simultaneously choose either "I'm Red" or "wait." If both participants choose "wait," then the match will continue to the next round. The match will end:

  (1) after round 5; or

  (2) after some round where there is at least one participant choosing "I'm Red."

  This round is called the "terminal round." Your payoff for this match depends on which round the terminal round is, your action in the terminal round, and your color. Important: your payoff does not depend on the other participant's color.

- Payoffs:

  (1) If you choose "wait" in the terminal round, you will get 0 points for this match regardless of your color.

  (2) If you choose "I'm Red" in the terminal round, your payoff for this match depends on which round the terminal round is and your own color. The payoffs are summarized in the following table. Notice that in each match, you and the other participant will face the same payoff table.

| Terminal Round | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Your payoff if your color is Red | $p_1$ | $p_2$ | $p_3$ | $p_4$ | $p_5$ |
| Your payoff if your color is White | $-w_1$ | $-w_2$ | $-w_3$ | $-w_4$ | $-w_5$ |

  (3) Examples:

    a. If you choose "I'm Red" in round 3, you will earn $p_3$ points if your color is Red and $-w_3$ if your color is White.

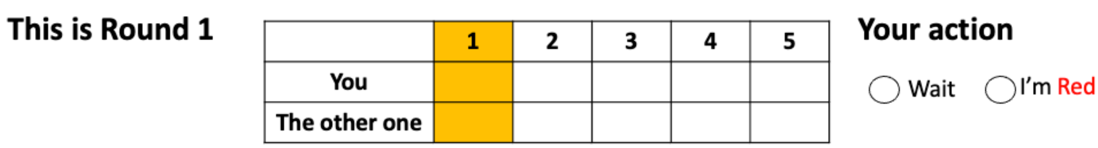
b. If you choose “wait” in round 4 and the other participant chooses “I’m Red” in the same round, you will get 0 points regardless of your color.

3. Decisions:

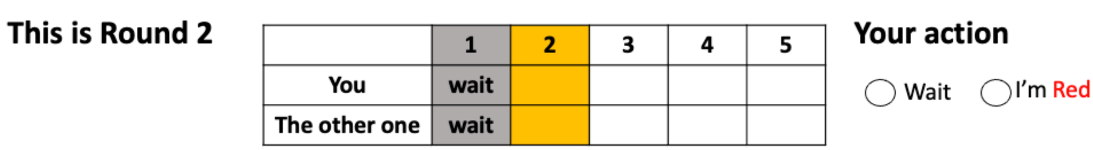
- After observing the other participant’s color, you and the other participant matched with you will play the game according to the rules described above.
- Therefore, your payoffs are summarized as below.

| | | You choose “I’m Red” in the terminal round | | You choose “wait” in the terminal round |
|---|---|---|---|---|
| Terminal Round | Your color | Red | White | Red or White |
| 1 | | $p_1$ | $-w_1$ | 0 |
| 2 | | $p_2$ | $-w_2$ | 0 |
| 3 | | $p_3$ | $-w_3$ | 0 |
| 4 | | $p_4$ | $-w_4$ | 0 |
| 5 | | $p_5$ | $-w_5$ | 0 |

- Each match starts from Round 1. You will make your decision in the following screen.




After you make your decision, the following would happen: If either you or the other participant chooses “I’m Red,” then this round is the terminal round, and your payoff is determined by your action in this round. However, if both you and the other participant choose “wait,” the match continues to the next round, and you will make your decision in the following screen.




Like the previous round, if either you or the other participant chooses “I’m Red,” the match will end after this round. Yet if both you and the other participant choose “wait,” the match will proceed to the next round.

- If the game proceeds to round 5, then the match will end after this round and your payoff is determined by your action (and color) in round 5.

4. At the end of each match, there will be a summary of the match which includes both of your colors, actions in each round (whenever applicable) and your own payoff for this match.

5. At the beginning of the experiment, you will start from 900 points. You will get paid in cash based on your total points earned from the 12 matches. If your total point is negative, you will only receive the show-up fee.

6. Important:

a. After each match, you will be randomly paired with another participant in the next match.

b. Your color and the other participant's color will also be randomly re-drawn in each match. The colors in each match are independent of the colors in other matches.

c. The probability distribution of colors and the payoff table will change in each match.

Please raise your hand if you have any questions. The question will be answered so that everyone can hear.

## 3.2 Simultaneous Treatment

### General Instructions

Thank you for participating in the experiment. You are about to take part in a decision-making experiment, in which your earnings will depend partly on your decisions, partly on the decision of others, and partly on chance.

The entire session will take place through computer terminals, and all interactions between participants will be conducted through the computers. Please do not talk or in any way try to communicate with other participants during the session.

The main task of the experiment consists of 12 matches. Before the main task, you will be asked to complete some comprehension questions. If you have any questions, please raise your hand and the question will be answered so that everyone can hear.

In this experiment, you will earn "points" in each match. Your earnings will be determined by the total points you earn in the 12 matches. Each point has a value of $0.02. That is, every 100 points generates $2 in earnings for you. In addition to your earnings from decisions, you will receive a show-up fee of $10. At the end of the experiment, your earnings will be rounded up to the nearest dollar amount. All your earnings will be paid in cash privately at the end of the experiment.

### Main Task

1. In this experiment, you will be asked to make decisions in 12 matches. You will be randomly matched with another participant into a group for every separate match. This random pairing changes in every match.

2. Each match in this experiment corresponds to a game with the following rules.

- At the beginning of each match, each of you and the other participant will be randomly assigned a "color" (either Red or White). After the colors are assigned, you will be able to see the color of the other participant who is paired with you. However, you cannot see your own color!
- There are 3 possible situations, and the probabilities of these situations are summarized in the following table.

| Situations | Probabilities |
|---|---|
| You are Red and the other participant is White. | $p$ |
| You are White and the other participant is Red. | $p$ |
| Both of you are Red. | $1-2p$ |

  In other words, there is always at least one Red participant among each group.

- Each match is played in rounds. In each match, there are at most 5 rounds. Your color and the other participant's color are fixed in the match. In each round, you and the other participant will simultaneously choose either "I'm Red" or "wait." If both participants choose "wait," then the match will continue to the next round. The match will end:

  (1) after round 5; or

  (2) after some round where there is at least one participant choosing "I'm Red."

  This round is called the "terminal round." Your payoff for this match depends on which round the terminal round is, your action in the terminal round, and your color. Important: your payoff does not depend on the other participant's color.

- Payoffs:

  (1) If you choose "wait" in the terminal round, you will get 0 points for this match regardless of your color.

  (2) If you choose "I'm Red" in the terminal round, your payoff for this match depends on which round the terminal round is and your own color. The payoffs are summarized in the following table. Notice that in each match, you and the other participant will face the same payoff table.

| Terminal Round | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Your payoff if your color is Red | $p_1$ | $p_2$ | $p_3$ | $p_4$ | $p_5$ |
| Your payoff if your color is White | $-w_1$ | $-w_2$ | $-w_3$ | $-w_4$ | $-w_5$ |

(3) Examples:

a. If you choose "I'm Red" in round 3, you will earn $p_3$ points if your color is Red and $-w_3$ if your color is White.

b. If you choose "wait" in round 4 and the other participant chooses "I'm Red" in the same round, you will get 0 points regardless of your color.

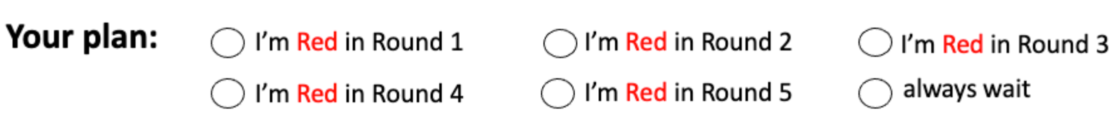
3. Decisions:

- Instead of playing the game round by round, after observing the other participant's color, you and the other participant are asked to simultaneously choose a "plan" which describes how you would commit to play the game if the game were played round by round. After you and the other participant both submit the plans, the computer will implement the plans and your payoff is determined accordingly.

- Since the game ends after some participant chooses "I'm Red," there are six possible plans corresponding to the earliest round you intend to choose "I'm Red" or "always wait." Specifically, the six plans are listed below.

  - **"I'm Red in Round 1"** means you plan to choose "I'm Red" in Round 1.
  - **"I'm Red in Round 2"** means you plan to choose "wait" in Round 1 and choose "I'm Red" in Round 2.
  - **"I'm Red in Round 3"** means you plan to choose "wait" in Round 1 and Round 2 and choose "I'm Red" in Round 3.
  - **"I'm Red in Round 4"** means you plan to choose "wait" in Round 1 to Round 3 and choose "I'm Red" in Round 4.
  - **"I'm Red in Round 5"** means you plan to choose "wait" in Round 1 to Round 4 and choose "I'm Red" in Round 5.
  - **"Always wait"** means you plan to choose "wait" in Round 1 to Round 5.

- In each match, you will be asked to choose your plan in the following screen.




- Therefore, your payoffs are summarized as below.

| | | You choose "I'm Red" **no later than** the other participant | | You choose "I'm Red" **later** or choose **"always wait"** |
|---|---|---|---|---|
| Terminal Round | Your color | Red | White | Red or White |
| 1 | | $p_1$ | $-w_1$ | 0 |
| 2 | | $p_2$ | $-w_2$ | 0 |
| 3 | | $p_3$ | $-w_3$ | 0 |
| 4 | | $p_4$ | $-w_4$ | 0 |
| 5 | | $p_5$ | $-w_5$ | 0 |

4. At the end of each match, there will be a summary of the match which includes both of your colors, the terminal round, your action, and your own payoff for this match. If you choose "I'm Red" later or at the same round as the other participant, you will be informed the other participant's exact plan. Otherwise, you will be told that the other participant is "later than you."

5. At the beginning of the experiment, you will start from 900 points. You will get paid in cash based on your total points earned from the 12 matches. If your total point is negative, you will only receive the show-up fee.

6. Important:

a. After each match, you will be randomly paired with another participant in the next match.

b. Your color and the other participant's color will also be randomly re-drawn in each match. The colors in each match are independent of the colors in other matches.

c. The probability distribution of colors and the payoff table will change in each match.

Please raise your hand if you have any questions. The question will be answered so that everyone can hear.

## 3.3 Screenshots

Figures 1 and 2 show the actual screenshots of the sequential treatment, and Figures 3 to 5 display the actual screenshots of the simultaneous treatment. Notice that Figure 4 represents the feedback screen of a player who selects "I'm Red" earlier than the opponent, and Figure 5 provides the perspective from the other player.

**Match 1**

| | You | The Other Participant |
|---|---|---|
| You see this | ?? | Red |

**Attention:**
1. There is at least one Red participant in this group.
2. The probabilities and the payoff table in each match are different.

| | Your color: Red<br>The other's: White | Your color: White<br>The other's: Red | Your color: Red<br>The other's: Red |
|---|---|---|---|
| Probability | 1/4 | 1/4 | 1/2 |

**Payoffs (in points):**

| | | You choose "I'm Red" in the terminal round | | You choose "wait" in the terminal round |
|---|---|---|---|---|
| Terminal Round | Your Color | Red | White | Red or White |
| 1 | | 100 | -444 | 0 |
| 2 | | 95 | -422 | 0 |
| 3 | | 90 | -401 | 0 |
| 4 | | 86 | -381 | 0 |
| 5 | | 81 | -362 | 0 |

**This is Round 1**

| | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| You | | | | | |
| The other one | | | | | |

**Your Action**

Wait I'm Red Confirm

Figure 1: The decision stage of the sequential treatment

**End of Match 1**

| | You | The Other Participant |
|---|---|---|
| The true colors | Red | Red |

**Payoffs (in points):**

| | | You choose "I'm Red" in the terminal round | | You choose "wait" in the terminal round |
|---|---|---|---|---|
| Terminal Round | Your Color | Red | White | Red or White |
| 1 | | 100 | -444 | 0 |
| 2 | | 95 | -422 | 0 |
| 3 | | 90 | -401 | 0 |
| 4 | | 86 | -381 | 0 |
| 5 | | 81 | -362 | 0 |

| Round | 1 | 2 | 3 | 4 | 5 | Colors |
|---|---|---|---|---|---|---|
| You | Wait | Wait | Red | -- | -- | Red |
| The other participant | Wait | Wait | Wait | -- | -- | Red |

Your payoff in this match: **90 points**

Confirm

Figure 2: The feedback stage of the sequential treatment

**Match 1**

| | You | The Other Participant |
|---|---|---|
| You see this | ?? | Red |

**Attention:**
1. There is at least one **Red** participant in this group.
2. The probabilities and the payoff table in each match are different.

| | Your color: **Red**<br>The other's: **White** | Your color: **White**<br>The other's: **Red** | Your color: **Red**<br>The other's: **Red** |
|---|---|---|---|
| Probability | 1/4 | 1/4 | 1/2 |

**Payoffs (in points):**

| | | You choose "I'm Red" **no later than** the other participant | | You choose "I'm Red" **later** or choose "**always wait**" |
|---|---|---|---|---|
| Terminal Round | Your Color | **Red** | **White** | **Red or White** |
| 1 | | 100 | -444 | 0 |
| 2 | | 95 | -422 | 0 |
| 3 | | 90 | -401 | 0 |
| 4 | | 86 | -381 | 0 |
| 5 | | 81 | -362 | 0 |

**Your plan:** ○ I'm Red in Round 1 ○ I'm Red in Round 2 ○ I'm Red in Round 3 ○ I'm Red in Round 4 ○ I'm Red in Round 5 ○ always wait

Confirm

Figure 3: The decision stage of the simultaneous treatment

**End of Match 1**

| | You | The Other Participant |
|---|---|---|
| The true colors | **Red** | **White** |

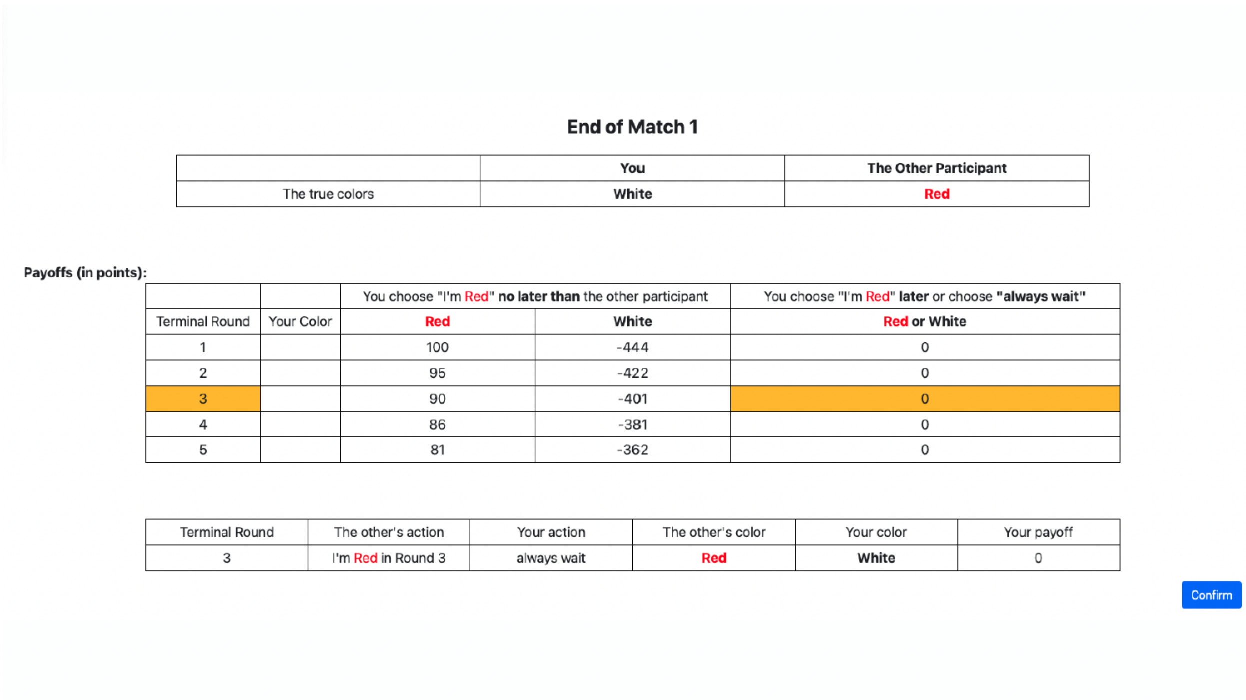
**Payoffs (in points):**

| | | You choose "I'm Red" **no later than** the other participant | | You choose "I'm Red" **later** or choose "**always wait**" |
|---|---|---|---|---|
| Terminal Round | Your Color | **Red** | **White** | **Red or White** |
| 1 | | 100 | -444 | 0 |
| 2 | | 95 | -422 | 0 |
| 3 | | 90 | -401 | 0 |
| 4 | | 86 | -381 | 0 |
| 5 | | 81 | -362 | 0 |

| Terminal Round | The other's action | Your action | The other's color | Your color | Your payoff |
|---|---|---|---|---|---|
| 3 | **Later than you** | I'm Red in Round 3 | **White** | **Red** | 90 |

Confirm

Figure 4: The feedback stage of the simultaneous treatment



**End of Match 1**

| | You | The Other Participant |
|---|---|---|
| The true colors | White | Red |

**Payoffs (in points):**

| | | You choose "I'm Red" **no later than** the other participant | | You choose "I'm Red" **later** or choose **"always wait"** |
|---|---|---|---|---|
| Terminal Round | Your Color | Red | White | Red or White |
| 1 | | 100 | -444 | 0 |
| 2 | | 95 | -422 | 0 |
| 3 | | 90 | -401 | 0 |
| 4 | | 86 | -381 | 0 |
| 5 | | 81 | -362 | 0 |

| Terminal Round | The other's action | Your action | The other's color | Your color | Your payoff |
|---|---|---|---|---|---|
| 3 | I'm Red in Round 3 | always wait | Red | White | 0 |



Figure 5: The feedback stage of the simultaneous treatment

# References

**Camerer, Colin F, Teck-Hua Ho, and Juin-Kuan Chong**, "A Cognitive Hierarchy Model of Games," *Quarterly Journal of Economics*, 2004, *119* (3), 861–898.

$a_i$ and information set $\mathcal{I}_i$:

$$\ln L^{\bar{D}}(\lambda, \tau) = \sum_{i} \sum_{\mathcal{I}_i \in \Pi_i} \sum_{a_i \in A_i(\mathcal{I}_i)} \mathbb{1}\{a_i, \mathcal{I}_i\} \ln \left[ \bar{\mathcal{D}}(a_i | \mathcal{I}_i, \lambda, \tau) \right],$$

where $\mathbb{1}\{a_i, \mathcal{I}_i\}$ is the indicator function which is 1 when subject $i$ chooses $a_i$ at $\mathcal{I}_i$.