Conditional Hypothesis Testing Systems under Non-Bayesian Updating*

Xiao Luo[a], Xuewen Qian[b]

[a]*Department of Economics, National University of Singapore, Singapore 117570*

[b]*School of Economics, University of Nottingham Ningbo China, Ningbo, 315100, China*

June 2022

Abstract

We introduce the notion of a Conditional Hypothesis Testing System (CHTS) that specifies the concrete mode of non-Bayesian reactions for unexpected news in a conditional world. We show that Ortoleva's (2012) Hypothesis Testing Model (HTM) gives rise to a CHTS and thus provides an axiomatic characterization of CHTS. Moreover, we show that the HTM updating rule is order independent—i.e., the order of receiving information does not influence the final posterior under the non-Bayesian updating rule. The notion of a CHTS enhances our understanding of HTM in a conditional world: It is formally linked to the conditional-probability-system expected utility model of Myerson (1986a). *JEL Classification: D81, D83.*

Keywords: HTM; CHTS; Dynamic coherence; Subjective expected utility; Non-Bayesian updating; Order independence

*Acknowledgement to be added. Financial support from the National University of Singapore is gratefully acknowledged.

1 Introduction

How a decision maker (henceforth DM) updates her beliefs after receiving new information is a foundational problem in information economics, game theory, and statistical theory. The standard way to model a DM's belief in a situation of uncertainty is by means of a prior—a subjective probability measure over states. Savage's (1954) subjective expected utility theory provide an axiomatic justification for the assumption that there exists a prior and the DM evaluates uncertain propsects or acts by the expected utility representation. Under subjective expected utility, applying Bayes' rule to the subjective probability is the standard way to update beliefs. That is, if a DM's belief is given by a measure π, and learns that an event A has occurred (i.e., $\pi(A) > 0$), then the DM's updating belief π_A—a probability measure conditional on A—is obtained by Bayes' formula:

$$\pi_A(A') = \frac{\pi(A \cap A')}{\pi(A)} \quad \text{for any arbitrary event } A'.$$

Throughout this paper, we write $\mathrm{BU}(\pi, A)$ for the Bayesian update of belief π, conditioning on non-zero probability event A. In a subjective expected utility framework, updating preferences by applying Bayes' rule to the subjective probability are dynamically consistent in a natural way that ex ante preferences are respected by updated preferences and, moreover, dynamic consistency implies such conditional measures must be the Bayesian updates (cf. Ghirardato (2002)).[1]

Experimental and empirical evidence, however, shows that people often systematically depart from Bayes' rule when confronted with new information (see, e.g., Benjamin (2019)). In an interesting paper, Ortoleva (2012) presents a Hypothesis Testing Model (henceforth HTM) to model the change in a paradigm in light of (non-)Bayesian reactions to

[1] Epstein and Le Breton (1993) discuss the general connection between dynamic consistency and Bayesian updating. They show that if (i) preferences are "based on beliefs" and (ii) admit dynamically consistent updating in response to new information, then Savage's (1954) axioms except the Sure Thing Principle guarantee that there exists a unique prior that represents beliefs and conditional beliefs are obtained using Bayes' rule.

(un)expected news/events. Ortoleva axiomatically characterizes HTM in which a decision maker acts like a standard Bayesian DM as long as the information she receives is assigned, according to a prior belief, a probability above a certain threshold (ϵ). She might react differently if she receives unexpected information that she assigned a small (no more than ϵ), and possibly equal to zero, probability. More specifically, when the DM changes the paradigm upon receiving information that is unlikely or unexpected according to her current prior, she considers a prior over priors (ρ), updates it using Bayes' rule for second-order priors, and selects a new prior that the updated prior over priors assigns the highest likelihood. HTM can be used to explain empirical evidence of the violations of Bayesian updating beliefs. For instance, investors might have a standard behavior in "business as usual" situations, but would be likely to change their beliefs in the case of unexpected circumstances. One important feature of this approach is that the HTM framework can accommodate non-Bayesian reactions to unexpected events, including zero-probability events, while the Bayesian paradigm is silent on what should happen on zero-probability events.

Ortoleva (2012, Theorem 1) shows that a novel behavioral postulate, called Dynamic Coherence (in lieu of Dynamic Consistency), together with other standard postulates, delivers a Hypothesis Testing representation (u, ρ, ϵ), where u is a utility function over consequences; ρ is a prior over priors on a state space; and ϵ is a threshold between 0 and 1. In particular, the (non-)Bayesian updating of a prior belief π_Ω, conditional on any arbitrary event A, can be represented by

$$\pi_A = \begin{cases} \mathrm{BU}(\pi_\Omega, A) & \text{if } \pi_\Omega(A) > \epsilon \\ \mathrm{BU}(\pi_A^*, A) & \text{otherwise} \end{cases}$$

where $\pi_\Omega = \pi_\Omega^*$ and $\{\pi_A^*\} = \arg\max_{\pi \in \Delta(\Omega)} \rho(\pi)\, \pi(A)$.

However, the HTM framework is presented in a parsimonious manner. While Ortoleva's (2012) Hypothesis Testing representation describes how to update a prior when an unexpected event occurs, it does not state the specific mode of new hypothesis testing, if the current one is rejected, which is an essential component of our framework in this paper. In economics, it is important and necessary to take into account zero-probability events and require well-defined conditional probabilities for them. For example, in game-theoretic contexts, such an omission is not as innocuous as it may seem: Indeed, if an event were assigned zero probability in a Nash equilibrium, then the question of what a player would do in this event is essential, because the equilibrium should specify strategic behavior for the player if this event occurred. In sequential games, Kreps and Wilson (1982) show that the detailed structure of a player's beliefs and preferences after he observes a zero-probability event may be crucial in the analysis of a game.[2] As Epstein and Le Breton (1993, p. 2) point out, "a satisfactory treatment of updating is a prerequisite for fruitful application of models of non-Bayesian beliefs or probabilistically non-sophisticated preferences, whether to intertemporal problems, game theory, or statistical theory." Our approach provides a thorough theory via the comprehensive specification of the (non-)Bayesian updating rule, including the specification of non-Bayesian reactions for unexpected news in unexpected circumstances.

In this paper, we introduce a novel notion of a Conditional Hypothesis Testing System (henceforth CHTS) that explicates the mode of non-Bayesian reactions for unexpected news in all hypothetical circumstances. We show that HTM gives rise to a CHTS representation $(u, \rho, \boldsymbol{\epsilon})$, where $\boldsymbol{\epsilon}= (\epsilon_A)_{A\in\Sigma}$ is a vector of thresholds conditioning on each and every contingent event $A \in \Sigma$ (see Theorem 2). In spite of violations of Bayes' rule, the mode of hypothesis testing beliefs in HTM appears to be the same mode in a conditional decision problem via updating. That is, the (non-)Bayesian belief updating in HTM conforms to a coherent structure: For any $A, A' \in \Sigma$ and $A \subseteq A'$,

$$\pi_A = \begin{cases} \mathrm{BU}(\pi_{A'}, A) & \text{if } \pi_{A'}(A) > \epsilon_{A'} \\ \mathrm{BU}(\pi^*_A, A) & \text{otherwise} \end{cases}$$

[2] Other notable examples, such as backward induction, forward induction, and iterated weak dominance in games, demonstrate the importance that players think about what would happen on zero-probability events; cf., e.g., Aumann (1995); Battigallia and Siniscalchi (2002); Samuelson (1992); Börgers (1994); and Brandenburger et al. (2008). As von Neumann and Morgenstern (1944) put it, "this is even more fundamental, the rules of rational behavior must provide definitely for the possibility of irrational conduct on the part of others ... description must include rules of conduct for all conceivable situations including those where 'the others' behaved irrationally, in the sense of the standards which the theory will set for them."

where $\pi_\Omega = \pi^*_\Omega$ and $\{\pi^*_A\} = \arg\max_{\pi\in\Delta(\Omega)} \rho(\pi)\,\pi(A)$. In other words, the new belief updating π_A conforms to the same mode of non-Bayesian reactions for unexpected news in a conditional world A' (instead of Ω). The threshold vector $\boldsymbol{\epsilon}$ here is interpreted as a confidence-level gauging system the DM adopts for "testing" her prior and "rejecting" the hypothesis in every contingent situation. The CHTS representation explicitly specifies the form of updating rule in all hypothetical circumstances, including those for which the DM receives unexpected information. As a result, we extend HTM to an enriched framework for conditional hypothesis testing decisions, which is important for the application of HTM in dynamic settings. When the threshold vector $\epsilon = \mathbf{0}$, our approach also provides a formal relation with the conditional-probability-system expected utility model of Myerson (1986a).

Our paper studies the (non-)Bayesian updating rule in the HTM framework of Ortoleva (2012) in which the DM is a standard expected utility maximizer. To allow for a non-Bayesian behavior, the HTM invokes the behavioral postulate of Dynamic Coherence in place of that of Dynamic Consistency. Despite their differences, Dynamic Coherence is intimately linked with a strong version of Dynamic Consistency—called "Conditional Dynamic Consistency" (henceforth CDC)—that requires Dynamic Consistency holds unwaveringly for all conditioning events including null events. Through CDC, the HTM framework provides a foundation for the conditional-probability-system expected utility model of Myerson (1986a) (see Corollary 1). Furthermore, we show that the HTM updating rule satisfies a natural form of order independence—that is, the order of receiving information never influences the final posterior under the HTM updating rule, as long as the information is not logically contradictory (see Theorem 3).

Our approach relates to the literature on non-Bayesian updating. A few recent papers study various models of non-Bayesian updating in a Bayesian environment. Epstein (2006) and Epstein, Noor, and Sandroni (2008) study a multi-period setting in which an agent does not update according to Bayes' rule: the agent might be tempted to use a posterior that is different from the Bayesian update of their prior. Zhao (2018; 2022) investigates the similarity-based and Pseudo-Bayesian updating rules in a framework in which information takes a qualitative form "event A is similar to event B" or "event A is more likely than event B." These papers do not deal with non-Bayesian updating rules on zero-probability events. By contrast, our paper allows to study non-Bayesian updating rules on zero-probability events, through the lens of CHTS, the main objective of this paper. In a closely related paper, Basu (2019) studies Bayesian updating rules and AGM belief revision, which can be used to define posteriors after all events, including zero probability events. By contrast, our paper focuses on non-Bayesian updating rules (see Section 5 for more discussion).

Deviations from Bayesian updating may also arise under ambiguity. There is a large literature on ambiguity. In the literature, ambiguity is usually modeled by relaxing Savage's (1954) Sure-Thing Principle while keeping other aspects of the standard Bayesian model intact. For instance, Gilboa and Schmeidler (1989) axiomatize the model of multiple-priors preferences that can accommodate ambiguity aversion; among others, Gilboa and Schmeidler (1993), Pires (2002), and Epstein and Schneider (2003) study updating rules for multiple priors. In the setting of dynamic choice under ambiguity, ambiguity-sensitive preferences must either relax Dynamic Consistency or other behavioral postulate(s); cf. Hanany and Klibanoff (2007); Siniscalchi (2009); and Galanis (2021) for a related discussion. The major difference in this paper is that our analysis of the (non-)Bayesian updating rule is conducted in the standard case that satisfies the Sure-Thing Principle, but allows for violations of Dynamic Consistency under the circumstances. As emphasized before, we focus mainly on violations of Bayes' rule in the HTM framework in which the DM is a Bayesian agent who is however open to fundamentally shifting her worldview.[3]

The rest of the paper is organized as follows. Section 2 provides an illustrative example, Section 3 briefly describes the HTM framework, notation, and definitions, Section 4 presents

[3]Dominiak et al. (2021) offer a systematic way–namely "minimum distance belief updating"—to extend Bayesian updating to general information and zero-probability events. They provide a behavioral characterization of "extended Bayesian updating" by a key axiom of "Informational Dynamic Consistency." They also show that the minimum-distance belief-updating approach is applicable to the HTM updating rule.

our CHTS Representation Theorem, Section 5 studies order independence of the HTM updating rule, and Section 6 concludes. All proofs are relegated to the Appendix.

2 An illustrative example

We provide an example to explain the main features in the frameworks of Ortoleva's (2012) HTM and our CHTS. We focus on the question: how to update beliefs upon the arrival of new information, with special attention to non-Bayesian reactions to (latent) information in hypothetical circumstances. Let $\Omega = \{\omega^1, \omega^2, \omega^3, \omega^4, \omega^5\}$ be a state space. Consider an HTM with a threshold $\epsilon = 0.02$ and a prior over priors $\rho \in \Delta(\Delta(\Omega))$ such that

$$\begin{cases} \rho = 0.70\pi^0 + 0.11\pi^1 + 0.10\pi^2 + 0.09\pi^3 \\ \pi^0 = 0.01\omega^2 + 0.01\omega^3 + 0.48\omega^4 + 0.50\omega^5 \\ \pi^1 = 0.10\omega^1 + 0.40\omega^2 + 0.50\omega^3 \\ \pi^2 = 1\omega^2 \\ \pi^3 = 1\omega^3 \end{cases}.$$

In this model, the DM selects the highest likelihood prior π^0. For any contingent event $A \in \Sigma$, she determines a conditional belief π_A using Bayes' rule whenever $\pi^0(A) > \epsilon$. However, if the information is "unexpected" (i.e., $\pi^0(A) \leq \epsilon$), the DM updates her prior over priors using Bayes' rule and picks the new maximum likelihood prior as her conditional belief π_A. More specifically,

$$\pi_A = \begin{cases} \mathrm{BU}(\pi^0, A) & \text{if } A \cap \{\omega^4, \omega^5\} \neq \emptyset \\ \pi^1 & \text{if } A = \{\omega^1, \omega^2, \omega^3\} \\ \pi^2 & \text{if } A = \{\omega^1, \omega^2\} \\ \pi^3 & \text{if } A = \{\omega^1, \omega^3\} \\ \mathrm{BU}(\pi^1, A) & \text{if } A = \{\omega^2, \omega^3\} \end{cases}$$

where $\mathrm{BU}(\bullet, A) \in \Delta(\Omega)$ denotes Bayesian update using A.

Now consider an "unexpected" event $A = \{\omega^1, \omega^2, \omega^3\}$. We identify a (minimal) threshold $\epsilon_A = 0.6$ that disciplines the reaction to all hypothetical events, conditioning on unexpected event A. Notably, the conditional belief π_A displays a consistent pattern of non-Bayesian reactions to hypothetical unexpected news by replacing $\epsilon = 0.02$ with $\epsilon_A = 0.6$. In the same vein, we spot a (minimal) threshold $\epsilon_A = 0.04$ for an "expected" event $A = \{\omega^1, \omega^2, \omega^3, \omega^4\}$. As a consequence, we have a vector of thresholds $\boldsymbol{\epsilon} = (\epsilon_A)_{A \in \Sigma}$ such that

$$\epsilon_A = \begin{cases} 0.02 & \text{if } A = \Omega \\ 0.04 & \text{if } A = \{\omega^1, \omega^2, \omega^3, \omega^4\} \\ 0.6 & \text{if } A = \{\omega^1, \omega^2, \omega^3\} \end{cases} .$$

That is, we extend the HTM framework to a CHTS framework. The main result of this paper shows that the existence of the CHTS representation $(u, \rho, \boldsymbol{\epsilon})$, where $\boldsymbol{\epsilon}= (\epsilon_A)_{A \in \Sigma}$, can be fully characterized by Dynamic Coherence in Ortoleva (2012), along with other standard behavioral postulates (see CHTS Representation Theorem in Section 4).

The (minimal) threshold vector $\boldsymbol{\epsilon} = (\epsilon_A)_{A \in \Sigma}$ in CHTS can be interpreted as a confidence-level gauging system the DM adopts for testing her prior and rejecting the hypothesis in all hypothetical scenarios—that is, threshold ϵ_A measures the degree of violation of Bayes' rule conditioning on any arbitrary event A. Observe that for a non-zero probability event $A = \{\omega^2, \omega^3\}$, π_A is not the Bayesian update of π^0, but rather the Bayesian update of π^1 because π_A =BU(π^1, A). Put another way, π_A is dynamically consistent with $\pi_{A'}$ (conditioning on $A' = \{\omega^1, \omega^2, \omega^3\}$), although it is not dynamically consistent with π_Ω (conditioning on Ω). The thresholds $\epsilon_\Omega = 0.02$ and $\epsilon_{A'} = 0.6$ signify distinct gauges for the violations of Bayes' rule conditional on Ω and A', respectively. We show that CHTS ($\boldsymbol{\epsilon}= \mathbf{0}$) is formally related to the conditional-probability-system expected utility model of Myerson (1986a) through a strong form of "Dynamic Consistency" (see Corollary 1 in Section 3). Our study sheds light on the dynamic structure of HTM.

3 The HTM framework: Notation and definitions

We adopt the notation and definitions of Ortoleva (2012) and list the necessary symbols, with very brief explanations, as follows:

Notation	Terminology
Ω	finite state space, with typical state $\omega \in \Omega$
X	(nonempty) set of consequences, with typical $x, y \in X$
$\Delta(\Omega)$	set of all probability measures (beliefs) on Ω, with typical $\pi \in \Delta(\Omega)$
$\Delta(\Delta(\Omega))$	set of all priors over priors on Ω, with typical $\rho \in \Delta(\Delta(\Omega))$
Σ	set of all nonempty subsets (events) of Ω, with typical $A \in \Sigma$
$\mathcal{F}$	set of all functions (acts) from Ω to X, with typical acts $f, g \in \mathcal{F}$
$\succeq_A$	a (complete, nondegenerate) preference relation over $\mathcal{F}$ conditional on A
fAg	$fAg \equiv \begin{cases} f(\omega) & \text{if } \omega \in A \\ g(\omega) & \text{if } \omega \in \Omega \backslash A \end{cases}$
$\mathrm{BU}(\pi, A)$	Bayesian update of belief π conditioning on A
$\mathrm{BU}(\rho, A)$	Bayesian update of a prior over priors ρ conditioning on A

AXIOM 1 (Well-Behaved Standard Preferences (WbP)): For any $A \in \Sigma$ and $f, g, h \in \mathcal{F}$:

(i) (Continuity): the sets $\{\alpha \in [0,1] : \alpha f + (1-\alpha) g \succeq_A h\}$ and $\{\alpha \in [0,1] : h \succeq_A \alpha f + (1-\alpha) g\}$ are closed.

(ii) (Independence): for any $\alpha \in (0,1)$

$$f \succeq_A g \Leftrightarrow \alpha f + (1-\alpha) h \succeq_A \alpha g + (1-\alpha) h.$$

(iii) (Monotonicity): if $f(\omega) \succeq_A g(\omega)$ for all $\omega \in \Omega$, then $f \succeq_A g$.

(iv) (Constant Preference Invariance): for any $x, y \in X$, $x \succeq_A y \Leftrightarrow x \succeq_\Omega y$.

WbP guarantees a standard expected utility representation (cf. Anscombe and Aumann (1963)): There exist a nonconstant affine function $u : X \to \mathbb{R}$ and $\pi_A \in \Delta(\Omega)$ for each $A \in \Sigma$ such that for any $f, g \in \mathcal{F}$

$$f \succeq_A g \Leftrightarrow \sum_{\omega \in \Omega} \pi_A(\omega) u(f(\omega)) \geq \sum_{\omega \in \Omega} \pi_A(\omega) u(g(\omega)).$$

AXIOM 2 (Consequentialism): For any $A \in \Sigma$, and $f, g \in \mathcal{F}$, if $f(\omega) = g(\omega)$ for all $\omega \in A$, then $f \sim_A g$.

Consequentialism is a basic requirement on conditional preferences that requires the complement of the conditioning event to be irrelevant for the conditional preference: If the DM is told that the true state lies inside some $A \in \Sigma$, then she is indifferent between two acts that differ only outside of A. That is, Consequentialism requires the complement of A to be $\succeq_A$-null.[4]

AXIOM 3 (Dynamic Consistency): For any $A \in \Sigma$, A not $\succeq_\Omega$-null, and for any $f, g \in \mathcal{F}$, we have

$$f \succeq_A g \Leftrightarrow fAg \succeq_\Omega g.$$

Dynamic Consistency requires that the arrival of some information A should not modify the ranking of two acts that coincide outside of A. Dynamic Consistency is the primary justification for Bayesian updating in the Bayesian model (cf. Ghirardato (2002)).

AXIOM 4 (Dynamic Coherence): For any $A_1, \ldots, A_n \in \Sigma$, if $(\Omega \backslash A_{i+1})$ is $\succeq_{A_i}$-null for $i = 1, \ldots, n-1$, and $(\Omega \backslash A_1)$ is $\succeq_{A_n}$-null, then $\succeq_{A_1} = \succeq_{A_n}$.

[4]We say A' is $\succeq_A$-null if $fA'g \sim_A g$ for any $f, g \in \mathcal{F}$. In the Bayesian framework, A' is not $\succeq_A$-null iff it is not a zero-probability event under the preference $\succeq_A$.

Dynamic Coherence captures an idea of "cycles in beliefs": If the informational content is the same in an information loop, then the DM's preferences should be the same after each piece of information—for example, in the Bayesian framework, the DM's posterior beliefs must be the same under the circumstances.

Definition 1: A class of preferences relations $\{\succeq_A\}_{A\in\Sigma}$ admits a *Hypothesis Testing (HT) representation* (u, ρ, ϵ) if there exist a nonconstant affine function $u : X \to \mathbb{R}$; a prior over priors $\rho \in \Delta(\Delta(\Omega))$ with finite-and-full support;[5] and a threshold $\epsilon \in [0, 1)$ such that for any $A \in \Sigma$, there exists $\pi_A \in \Delta(\Omega)$ such that

(i) for any $f, g \in \mathcal{F}$,

$$f \succeq_A g \Leftrightarrow \sum_{\omega\in\Omega} \pi_A(\omega) u(f(\omega)) \geq \sum_{\omega\in\Omega} \pi_A(\omega) u(g(\omega));$$

(ii) $\{\pi^*_\Omega\} = \arg\max_{\pi\in\Delta(\Omega)} \rho(\pi)$.

(iii)

$$\pi_A = \begin{cases} \mathrm{BU}(\pi_\Omega, A) & \text{if } \pi_\Omega(A) > \epsilon \\ \mathrm{BU}(\pi^*_A, A) & \text{if } \pi_\Omega(A) \leq \epsilon \end{cases}$$

where $\{\pi^*_A\} = \arg\max_{\pi\in\Delta(\Omega)} \mathrm{BU}(\rho, A)(\pi)$.[6]

In Definition 1, an HT representation (u, ρ, ϵ) is said to be *minimal* if there is no $0 \leq \epsilon' < \epsilon$ such that (u, ρ, ϵ') is an HT representation of the same preferences $\{\succeq_A\}_{A\in\Sigma}$. Ortoleva (2012) shows the following Hypothesis Testing representation theorem.

Theorem 1 (Ortoleva's (2012) HT Representation Theorem): *A class of preference relations $\{\succeq_A\}_{A\in\Sigma}$ satisfies Well-Behaved Standard Preferences (WbP), Consequentialism,*

[5] ρ has *finite-and-full support* if (i) $\mathrm{supp}(\rho)$ is finite; and (ii) $\Omega = \cup_{\pi\in\mathrm{supp}(\rho)}\mathrm{supp}(\pi)$.

[6] $\mathrm{BU}(\rho, A) \in \Delta(\Delta(\Omega))$ is defined as $\mathrm{BU}(\rho, A)(\pi) = \frac{\pi(A)\rho(\pi)}{\sum_{\pi'\in\mathrm{supp}(\rho)} \pi'(A)\rho(\pi)}$ $\forall \pi \in \Delta(\Omega)$, provided $\pi'(A) > 0$ for some $\pi' \in \mathrm{supp}(\rho)$.

and Dynamic Coherence if and only if it admits a minimal HT representation (u, ρ, ϵ)*. Moreover,* $\epsilon = 0$ *if and only if* $\{\succeq_A\}_{A\in\Sigma}$ *also satisfies Dynamic Consistency.*

4 CHTS Representation Theorem

We introduce the notion of a Conditional Hypothesis Testing System.

Definition 2: A class of preferences relations $\{\succeq_A\}_{A\in\Sigma}$ admits a *Conditional Hypothesis Testing System (CHTS) representation* $(u, \rho, \boldsymbol{\epsilon})$ if there exist a nonconstant affine function $u : X \to \mathbb{R}$; a prior over priors $\rho \in \Delta(\Delta(\Omega))$ with finite-and-full support; and a vector of thresholds $\boldsymbol{\epsilon}= (\epsilon_A)_{A\in\Sigma}$ (where $\epsilon_A \in [0, 1)$) such that for any $A \in \Sigma$, there exists $\pi_A \in \Delta(\Omega)$ such that

(i) for any $f, g \in \mathcal{F}$,

$$f \succeq_A g \Leftrightarrow \sum_{\omega\in\Omega} \pi_A(\omega)\, u(f(\omega)) \geq \sum_{\omega\in\Omega} \pi_A(\omega)\, u(g(\omega))\,;$$

(ii) for any $A' \in \Sigma$ with $A' \supseteq A$,

$$\pi_A = \begin{cases} \mathrm{BU}(\pi_{A'}, A) & \text{if } \pi_{A'}(A) > \epsilon_{A'} \\ \mathrm{BU}(\pi^*_A, A) & \text{if } \pi_{A'}(A) \leq \epsilon_{A'} \end{cases}$$

where $\pi_\Omega = \pi^*_\Omega$ and $\{\pi^*_A\} = \arg\max_{\pi\in\Delta(\Omega)} \mathrm{BU}(\rho, A)(\pi)$.

We call such a class $\{\pi_A\}_{A\in\Sigma}$ a *Hypothesis-Testing Conditional Probability System (HTCPS).*

In Definition 2, a CHTS representation $(u, \rho, \boldsymbol{\epsilon})$ is said to be *minimal* if there is no $\mathbf{0} \leq \boldsymbol{\epsilon}' \lneqq \boldsymbol{\epsilon}$ such that $(u, \rho, \boldsymbol{\epsilon}')$ is a CHTS representation of the same preferences $\{\succeq_A\}_{A\in\Sigma}$. Obviously, if Condition (ii) is required only for $A' = \Omega$, then a CHTS representation $(u, \rho, \boldsymbol{\epsilon})$ is essentially a Hypothesis Testing representation (u, ρ, ϵ) in Ortoleva (2012).

For the purpose of this paper, we introduce a conditional version of Dynamic Consistency as follows.

AXIOM 3* (Conditional Dynamic Consistency (CDC)): For any $A, A' \in \Sigma$ and for any $f, g \in \mathcal{F}$, if $A \subseteq A'$ is not $\succeq_{A'}$-null, we have

$$f \succeq_A g \Leftrightarrow fAg \succeq_{A'} g.$$

CDC requires that the arrival of new information A should not change a preference $\succeq_{A'}$ over two acts that coincide outside of A, provided that $A \subseteq A'$ is not $\succeq_{A'}$-null. By contrast, Dynamic Consistency requires that the property holds merely for the ex ante preference relation $\succeq_\Omega$, rather than all preference relations $\succeq_{A'}$. In the subjective expected utility model, Dynamic Consistency requires that Bayes' rule holds for nonzero-probability events: It allows for violations of Bayes' rule for zero-probability events, whereas CDC requires full Bayesian behavior in the sense that Bayes' rule holds in all conditioning events, including zero-probability events. Evidently, it is a stronger form of Dynamic Consistency.

Dynamic Coherence (Axiom 4) is a key behavioral postulate for HTM. Dynamic Coherence appears to be a fairly natural condition: It looks weaker than Dynamic Consistency in the sense that the latter allows for violations of Dynamic Consistency in small-probability events or null events w.r.t. the ex ante preference relation. Our CHTS representation theorem reveals that Dynamic Coherence has an inherently consistent characteristic attribute: In every contingent event, Dynamic Coherence is indeed weaker than Dynamic Consistency in the same manner—that is, CDC implies Dynamic Coherence in HTM.[7]

[7] Ortoleva (2012, p. 2419) states "Dynamic Coherence is neither stronger nor weaker than Dynamic Consistency: while it is does allow for violations of Dynamic Consistency, albeit regulating them, it also disciplines the reaction to null events, on which Dynamic Consistency has no bite." Lemma (ii) in the Appendix asserts that Dynamic Coherence is unequivocally weaker than CDC under Consequentialism.

Theorem 2 (CHTS Representation Theorem): *A class of preference relations* $\{\succeq_A\}_{A\in\Sigma}$ *satisfies Well-Behaved Standard Preferences (WbP), Consequentialism, and Dynamic Coherence if and only if it admits a minimal CHTS representation* $(u, \rho, \boldsymbol{\epsilon})$*. Moreover,* $\boldsymbol{\epsilon} = 0$ *if and only if* $\{\succeq_A\}_{A\in\Sigma}$ *also satisfies Conditional Dynamic Consistency (CDC).*

Once we realize that the behavioral axioms of WbP, Consequentialism, and Dynamic Coherence are valid for any subclass of conditional preference relations (see Lemma 1 in the Appendix), then for any $A \in \Sigma$ there is a Hypothesis Testing representation (u, ρ, ϵ_A) for $\{\succeq_{A'}\}_{A'\in\Sigma_A}$, where Σ_A denotes the set of nonempty subsets of A. As a result, CHTS Representation Theorem follows from Ortoleva's (2012) proof.

The CHTS representation manifests that HTM displays a consistent pattern of behavior even if the DM has non-Bayesian reactions to unexpected news. From a normative point of view, this feature of HTM is desirable in analyzing complex models, possibly with non-Bayesian updating; an analyst can technically reduce the framework to a simple one by representing and analyzing the DM's dynamic choice problem as that of choosing a contingent act in a consistent manner. The second part of CHTS Representation Theorem shows that if Dynamic Coherence is further strengthened by CDC, the DM follows Bayes' rule whenever it applies in every conditioning event, but reconsiders which prior to use whenever she faces zero-conditional-probability events. Corollary 1 below makes a connection between HTM and the conditional-probability-system expected utility model in Myerson (1986a). Consequently, our approach provides an alternative foundation for the conditional-probability-system expected utility model.

Definition 3: A collection of probability measures $\{\pi_A\}_{A\in\Sigma}$ over Ω is a *conditional*

This result offers additional insight into why Dynamic Coherence is not enough to guarantee the full Bayesian behavior, despite the fact that it "imposes a coherence between beliefs across different pieces of information, independently of the belief before information" in a Bayesian framework.

probability system (CPS) if for any $A \in \Sigma$, (i) $\pi_A(A) = 1$ and (ii) [Bayes' formula] $\pi_B(C) = \pi_B(A)\,\pi_A(C)\ \forall C \subseteq A \subseteq B \subseteq \Omega$.

Corollary 1: *A class of preference relations $\{\succeq_A\}_{A\in\Sigma}$ satisfies WbP, Consequentialism, and CDC if and only if it admits a conditional-probability-system expected utility representation: There exist a nonconstant affine function $u : X \to \mathbb{R}$ and a CPS $\{\pi_A\}_{A\in\Sigma}$ such that for any $f, g \in \mathcal{F}$*

$$f \succeq_A g \Leftrightarrow \sum_{\omega\in\Omega} \pi_A(\omega)\, u(f(\omega)) \geq \sum_{\omega\in\Omega} \pi_A(\omega)\, u(g(\omega)).$$

Myerson (1986b) shows that any CPS can be characterized by the limit of a sequence of conditional probability systems generated by full-support probability distributions over Ω. The notion of CHTS offers an alternative characterization of CPS by disciplining Hypothesis Testing updating in each and every conditioning event.

Corollary 2: *An HTCPS $\{\pi_A\}_{A\in\Sigma}$ is a CPS on Ω iff there exists a "partitional" prior over priors ρ in the underlying CHTS representation—i.e., $\{\operatorname{supp}(\pi)\}_{\pi\in\operatorname{supp}(\rho)}$ is a partition of Ω—such that for any pair $A, A' \in \Sigma$ and $A \subseteq A'$,*

$$\pi_A = \begin{cases} \mathrm{BU}(\pi_{A'}, A) & \text{if } \pi_{A'}(A) > 0 \\ \mathrm{BU}(\pi_A^*, A) & \text{if } \pi_{A'}(A) = 0 \end{cases}$$

where $\pi_\Omega = \pi_\Omega^*$ and $\{\pi_A^*\} = \arg\max_{\pi\in\Delta(\Omega)} \rho(\pi)\,\pi(A)$.

Remark 1. The induced "partitional" priors $(\pi)_{\pi\in\operatorname{supp}(\rho)}$ in Corollary 2 can be viewed as a lexicographic conditional probability system (LCPS) in Blume et al. (1991) (cf. also Halpern (2007) for related discussions). Thus, a CHTS is also related to an LCPS.

Ortoleva (2012) shows that, if add Dynamic Consistency to the Hypothesis Testing Model, we obtain a characterization of a Bayesian model, in which the DM not only follows

Bayes' rule when it applies, but also the DM's beliefs are disciplined by Dynamic Coherence when Bayes' rule does not apply in null events. Our CHTS Representation Theorem immediately implies the following.

Corollary 3: *Consider a class of preference relations $\{\succeq_A\}_{A\in\Sigma}$. The following statements are equivalent:*

(1) *The class $\{\succeq_A\}_{A\in\Sigma}$ satisfies WbP, Consequentialism, Dynamic Coherence, and Dynamic Consistency.*

(2) *The class $\{\succeq_A\}_{A\in\Sigma}$ admits a Hypothesis Testing representation (u, ρ, ϵ) with $\epsilon = 0$.*

(3) *The class $\{\succeq_A\}_{A\in\Sigma}$ admits a CHTS representation $(u, \rho, \boldsymbol{\epsilon})$ that satisfies $\epsilon_\Omega = 0$.*

5 The HTM updating rule: Order independence

Our analytical framework can be used for the study of sequential information processing in a dynamic setting. In the actual situation, the belief updating problem is a recurring one of revising the current belief to a new belief after arrival of new information in the sequel. A good update rule should satisfy a "non-manipulability" property—order independence—that the order of receiving information does not change the final posterior belief (cf. Basu (2019); Sadler (2021)). Evidently, when restricted to a prior positive probability events, the Bayesian updating rule does satisfy this desirable property. Observe that Consequentialism implies fact-based updating that, in turn, implies the violation of invariant if the receiving information were to be contradictory (e.g., a pair of disjoint events); thus, order independence should be required only for the uncontradictory information under Consequentialism. Basu (2019, Claim 1) shows that no Bayesian updating rules satisfy a so-called "strong path independence" when the cardinality of state space is greater than two. He provides a complete characterization of a class of lexicographic updating rules by a weak form of path independence. Sadler (2021) studies a practical guide to updating beliefs from contradictory evidence in complex information environments. Sadler provides a general framework of belief formation, along with several updating axioms, that helps us identify tradeoffs inherent in the way we form beliefs.[8]

In this section, we show an order independence result about the HTM updating rule. More specifically, we define the HTM updating rule as a function that determines, conditioning on every event, a new belief for a prior belief and new information, regardless of unexpected or contradictory information. We show that the HTM updating rule satisfies the property that the update rule is invariant with respect to the order in which the new information arrives, provided the information is not logically contradictory.

An important feature of HTM is the decomposition of preferences into tastes and beliefs like the one in the standard Bayesian model. Consider an HTM that has a CHTS representation $(u, \rho, \boldsymbol{\epsilon})$, where $\boldsymbol{\epsilon} = (\epsilon_A)_{A \in \Sigma}$, with the associated HTCPS $\boldsymbol{\pi} = \{\pi_A\}_{A \in \Sigma}$. The belief part of the HTM can be represented by a belief-HTM $(\rho, \boldsymbol{\epsilon}, \boldsymbol{\pi})$.

Definition 4: An *updating rule* $\boldsymbol{\varphi} = \{\varphi_A\}_{A \in \Sigma}$ on state space Ω is defined as follow: For any event $A \in \Sigma$,

$$\varphi_A : \Delta(\Omega) \times \Sigma \to \Delta(\Omega)$$

[8] When information is qualitative, Zhao (2022) provides an axiom of Exchangeability, a variant form of order independence, which requires that the order in which the information arrives does not matter if the different pieces of information neither reinforce nor contradict each other. Zhao axiomatizes a class of Pseudo-Bayesian updating rules by Exchangeability, together with other axioms. In a learning context of recommendations, Ke, Wu, and Zhao (2022) study the contraction rule for belief updating, which is sensitive to the order in which recommendations arrive (e.g., recency bias).

such that for all $\pi \in \Delta(\Omega)$ and $\mathcal{A} \in \Sigma$,

$$\varphi_A(\pi, \mathcal{A}) \in \Delta(\mathcal{A}),$$

where $\Delta(\mathcal{A}) = \{\pi \in \Delta(\Omega) : \pi(\mathcal{A}) = 1\}$.

An updating rule $\boldsymbol{\varphi}$ is an *HTM updating rule* for a belief-HTM $(\rho, \boldsymbol{\epsilon}, \boldsymbol{\pi})$ if for any $A, \mathcal{A} \in \Sigma$ and $\pi \in \Delta(A)$,

$$\varphi_A(\pi, \mathcal{A}) = \begin{cases} \mathrm{BU}(\pi, \mathcal{A}) & \text{if } \pi(\mathcal{A}) > \epsilon_A \\ \pi_{A \cap \mathcal{A}} & \text{if } \pi(\mathcal{A}) \leq \epsilon_A \text{ and } A \cap \mathcal{A} \neq \emptyset \\ \pi_{\mathcal{A}} & \text{if } \pi(\mathcal{A}) \leq \epsilon_A \text{ and } A \cap \mathcal{A} = \emptyset \end{cases}.$$

Theorem 3 below shows that the updating rule adopted in HTM is insentive to the order in which the uncontradictory information appears—that is, conditioning on any arbitrary event A, for all consistent pairs of events $B, C \in \Sigma_A$ (i.e., $B \cap C \neq \emptyset$), the order in which information arrives never influences the final posterior under the updating rule. In other words, under the HTM updating rule, the order in which the information arrives does not matter if the different pieces of information do not contradict each other.

Theorem 3 (Order Independence): *The HTM updating rule, $\boldsymbol{\varphi} = \{\varphi_A\}_{A \in \Sigma}$, for a belief-HTM $(\rho, \boldsymbol{\epsilon}, \boldsymbol{\pi})$ satisfies order independence: For all $A \in \Sigma$ and $B, C \in \Sigma_A$ with $B \cap C \neq \emptyset$,*

$$\varphi_B(\varphi_A(\pi_A, B), C) = \varphi_A(\pi_A, B \cap C) = \varphi_C(\varphi_A(\pi_A, C), B).$$

Moreover, for all $\pi \in \Delta(\Omega)$ such that $\pi = \pi_\Omega$ in a belief-HTM $(\rho, \boldsymbol{\epsilon}, \boldsymbol{\pi})$, it is true that

$$\varphi_B(\varphi_\Omega(\pi, B), C) = \pi_{B \cap C} = \varphi_C(\varphi_\Omega(\pi, C), B),$$

for all $B, C \in \Sigma$ with $B \cap C \neq \emptyset$.

The order independence in Theorem 3 refers to a very general form of order independence that is applicable to (non-)Bayesian updating. Thus, the HTM updating rule that satisfies this order independence could violate Bayes' rule like other (non-Bayesian) statistical techniques. Corollary 4 below states that if the information is sufficiently correlated, order independence pertains to Bayesian updating.

Corollary 4: *The HTM updating rule, $\boldsymbol{\varphi} = \{\varphi_A\}_{A\in\Sigma}$, for $(\rho, \boldsymbol{\epsilon}, \boldsymbol{\pi})$ satisfies a BU-form of order independence: For all $A \in \Sigma$ and $B, C \in \Sigma_A$ with $\pi_A (B \cap C) > \epsilon_A$,*

$$\varphi_B (\varphi_A (\pi_A, B), C) = \mathrm{BU} (\pi_A, B \cap C) = \varphi_C (\varphi_A (\pi_A, C), B).$$

Remark 2. Due to the feature of fact-based updating (i.e., $\varphi_A (\pi, \mathcal{A}) \in \Delta (\mathcal{A})$), Theorem 3 implies that the HTM updating rule must satisfy order independence if, and only if, the information is not logically contradictory (i.e., $A \cap B \neq \emptyset$). Corollary 4 asserts that if the information is $\boldsymbol{\epsilon}$-sufficiently correlated (i.e., $\pi_A (B \cap C) > \epsilon_A$), the HTM updating rule turns out to be Bayesian updating. If $\boldsymbol{\epsilon} = \mathbf{0}$, Corollaries 1 and 4 yield an order independence result for conditional probability systems. That is, for any CPS $\{\pi_A\}_{A\in\Sigma}$ and for all correlation-consistent pairs of events $B, C \in \Sigma_A$ (i.e., $\pi_A (B \cap C) > 0$), we have $\varphi_B (\varphi_A (\pi_A, B), C) =\mathrm{BU}(\pi_A, B \cap C) = \varphi_C (\varphi_A (\pi_A, C), B)$. This is harmonious with Basu's (2019) result on (weak) path independence of lexicographic updating rule, because a CPS can alternatively be viewed as a "lexicographic conditional probability system." We would like to point out that our order independence theorem applies to a wider range of (non-)Bayesian updating rules, irrespective of whether the occurred events are expected or unexpected in a conditional world.

We end this section by providing an example to explain our order independence result about the HTM updating rule.

Example 1. Suppose $\Omega = \{\omega^1, \omega^2, \omega^3\}$ is a 3-state space. Consider an HTM with a threshold vector $\boldsymbol{\epsilon} = \mathbf{0}$ and a prior over priors $\rho \in \Delta(\Delta(\Omega))$ such that

$$\begin{cases} \rho = 0.6\pi^0 + 0.4\pi^1 \\ \pi^0 = 1\omega^1 \\ \pi^1 = 0.5\omega^2 + 0.5\omega^3 \end{cases}.$$

The associated HTCPS $\boldsymbol{\pi} = \{\pi_A\}_{A\in\Sigma}$ is a CPS on Ω such that

$$\pi_A = \begin{cases} 1\omega^1 & \text{if } A = \Omega, B \\ 0.5\omega^2 + 0.5\omega^3 & \text{if } A = C \end{cases},$$

where $B = \{\omega^1, \omega^2\}$ and $C = \{\omega^2, \omega^3\}$. Obviously, $B \cap C = \{\omega^2\} \neq \emptyset$. We have

$$\varphi_B(\varphi_\Omega(\pi_\Omega, B), C) = 1\omega^2 = \varphi_C(\varphi_\Omega(\pi_\Omega, C), B).$$

That is, this is an instance of our order independence theorem when $A = \Omega$, $B = \{\omega^1, \omega^2\}$ and $C = \{\omega^2, \omega^3\}$. Remarkably, in spite of $\pi_\Omega(C) = 0$, the HTM updating rule $\boldsymbol{\varphi}$ satisfies order independence under the circumstances.

Nevertheless, there is no Bayesian updating rule that satisfies "strong path independence" in Basu's (2019) sense. To explain this point, we let $\varphi^{\text{Basu}} : \Delta(\Omega) \times \Sigma \to \Delta(\Omega)$ denote a Bayesian updating rule for this CPS $\boldsymbol{\pi} = \{\pi_A\}_{A\in\Sigma}$ such that $\varphi^{\text{Basu}}(\pi_\Omega, C) = \pi_C$. Then,

$$\begin{cases} \varphi^{\text{Basu}}\left(\varphi^{\text{Basu}}(\pi_\Omega, B), C\right) = \varphi^{\text{Basu}}(1\omega^1, C) = 0.5\omega^2 + 0.5\omega^3; \\ \varphi^{\text{Basu}}\left(\varphi^{\text{Basu}}(\pi_\Omega, C), B\right) = \varphi^{\text{Basu}}(0.5\omega^2 + 0.5\omega^3, B) = 1\omega^2. \end{cases}$$

That is, the Bayesian updating rule φ^{Basu} fails to satisfy "strong path independence." The thrust of our order independence theorem is that the HTM updating rule, $\boldsymbol{\varphi} = \{\varphi_A\}_{A\in\Sigma}$, allows to define, conditioning on any $A \in \Sigma$, a new updated belief for a prior belief and new information, while φ^{Basu} does not. (Note that the HTM updating rule $\boldsymbol{\varphi}$ allows to have $\varphi_\Omega(\pi_\Omega, C) \neq \varphi_B(\pi_\Omega, C)$ for zero-probability event C.) Basu (2019) takes a different way to show (weak) path independence of lexicographic updating rule, which is consistent with our order independence result (cf. Remark 2).

6 Conclusion

Bayesian assessing every contingent event is infeasible in the practical application. We are often faced with the question of how to update beliefs in response to unexpected news. Ortoleva (2012) provides a decision-theoretic model of HTM in which an agent is Bayesian, but after observing an event that is sufficiently unlikely, she considers a prior over priors, selecting the prior that assigned the highest likelihood to the unexpected observation.

The main purpose of this paper is to extend the HTM framework to the one in a conditional world. In doing so, we introduce the notion of a Conditional Hypothesis Testing System (CHTS) that delineates the mode of non-Bayesian reactions for unexpected news in a conditional decision problem. We show that HTM gives rise to a CHTS representation $(u, \rho, \boldsymbol{\epsilon})$ by an array of thresholds $\boldsymbol{\epsilon}= (\epsilon_A)_{A\in\Sigma}$ (Theorem 2). More specifically, under the behavioral postulates—Well-Behaved Standard Preferences (WbP), Consequentialism, and Dynamic Coherence—HTM determines a CHTS representation that specifies the concrete mode of hypothesis testing in every conditioning event, including the circumstances under which the DM receives unexpected information that some rare or null event has occurred. As a result, our CHTS representation implies that the DM has an HTM theory in a recursive manner: The behavioral pattern of non-Bayesian reactions by applying the theory ex ante in anticipation of future contingencies do not contradict the behavioral pattern prescribed based on the theory via updating. The recursive structure is obviously desirable in computation and analysis of intertemporal problems. While the original HTM framework offers the key insight of hypothesis testing, the dynamic feature of CHTS is suppressed in Ortoleva's (2012) HT Representation Theorem. As we have emphasized, the specific form of hypothesis testing for the DM's reaction to unexpected information conditional on unexpected information is our main focus in this paper.

In this paper, we also demonstrate that our framework can be used for the study of sequential information processing in a dynamic setting. We show that the HTM updating rule satisfies order independence—i.e., the order of receiving uncontradictory information never influences the final posterior under the updating rule, regardless of which event has occurred (Theorem 3). If moreover the information is sufficiently correlated, the form of order independence turns out to be the standard one for Bayesian updating (Corollary 4). On a conceptual level, the notion of CHTS helps us better understand the HTM framework in a conditional world; for instance, it improves our understanding of the relation between HTM and other models such as the conditional-probability-system expected utility model (Corollary 1).[9] Like the concept of Conditional Probability System, the CHTS notion captures dynamically coherent characteristics whenever (latent) information becomes available in a hypothesis testing model: It can be useful for a systematic and rigorous analysis of various economic models of asymmetric information (e.g., Bayesian persuasion, sequential psychological games, and incomplete-information games), possibly played by non-Bayesian agents who may change the paradigm upon receiving new information. In this regard, CHTS provides a general analytical framework for studying complex situations that involve the hypothetical and counterfactual reasoning of Bayesian and non-Bayesian agents, by unlocking the inherently dynamic characteristics of HTM (whereby the DM envisions the possibility of receiving information in all hypothetical scenarios). In particular, CHTS offers a formal apparatus for inquiring into higher-order issues regarding "unexpected" expected events and expected "unexpected" events in future contingencies such as nuclear war, pandemics, global warming, etc.

[9] Karni and Vierø (2013, 2017) study belief updating in the wake of growing awareness. They characterize behaviorally "reverse Bayesianism" according to which the relatively likelihood ratio for events that the decision is already aware of remains unchanged upon becoming aware of a novel event; cf. also Schipper (2022) for a related study of "reverse" belief updating under subjective expected utility. The extension of our analysis to "reverse" (non-)Bayesian updating is an intriguing subject for further research.

7 Appendix: Proofs

Lemma 1: *(i) A class of preference relations $\{\succeq_A\}_{A\in\Sigma}$ satisfies WbP, Consequentialism, and Dynamic Coherence if and only if for any $A \in \Sigma$, the subclass of preference relations $\{\succeq_{A'}\}_{A'\in\Sigma_A}$ satisfies WbP, Consequentialism, and Dynamic Coherence. (ii) Under Consequentialism, CDC implies Dynamic Coherence.*

Proof: (i) The "If" part is trivial. Suppose $\{\succeq_A\}_{A\in\Sigma}$ satisfies WbP, Consequentialism, and Dynamic Coherence. Then, WbP and Consequentialism are also applied to every subclass $\{\succeq_{A'}\}_{A'\in\Sigma_A}$. Because "$(\Omega\backslash A_{i+1})$ is $\succeq_{A_i}$-null" implies "$(A_i\backslash A_{i+1})$ is $\succeq_{A_i}$-null" (where $i = 1, \ldots, n$, $A_i \in \Sigma_A$, and $A_{n+1} = A_1$), Dynamic Coherence holds true for the subclass $\{\succeq_{A'}\}_{A'\in\Sigma_A}$.

(ii) Let $i = 1, \ldots, n$, and let $A_i \in \Sigma$. Assume $\Omega\backslash A_{i+1}$ is $\succeq_{A_i}$-null (where $A_{n+1} = A_1$). Then, $A_i\backslash A_{i+1}$ is $\succeq_{A_i}$-null. Define $A^* = \cup_{i=1}^n A_i$. By CDC, $A_i\backslash A_{i+1}$ is $\succeq_{A^*}$-null; hence, $\cup_{i=1}^n (A_i\backslash A_{i+1})$ is $\succeq_{A^*}$-null. Note that $A^*\backslash [\cap_{i=1}^n A_i] = \cup_{i=1}^n (A_i\backslash A_{i+1})$. By Consequentialism, $\Omega\backslash [\cap_{i=1}^n A_i]$ is $\succeq_{A^*}$-null. Therefore, for $i = 1, \ldots, n$ and any $f, g \in \mathcal{F}$,

$$fA_ig \sim_{A^*} f \; ; \tag{1}$$

and by CDC,

$$f \succeq_{A_i} g \Leftrightarrow fA_ig \succeq_{A^*} g. \tag{2}$$

By (1) and (2), $f \succeq_{A_i} g \Leftrightarrow f \succeq_{A^*} g$ for any $f, g \in \mathcal{F}$; that is, $\succeq_{A_i}=\succeq_{A^*}$for $i = 1, \ldots, n$. Hence, Dynamic Coherence holds. $\square$

Proof of Theorem 2: Let $A \in \Sigma$. By Lemma 1(i) and Ortoleva's (2012) Theorem 1, there exists a minimal Hypothesis Testing representation (u, ρ_A, ϵ_A) for $\{\succeq_{A'}\}_{A'\in\Sigma_A}$. Again by Ortoleva's (2012) Theorem 1, there exists a minimal Hypothesis Testing representation $(u, \rho, \epsilon) = (u, \rho_\Omega, \epsilon_\Omega)$ for $\{\succeq_A\}_{A\in\Sigma}$. Since the prior over priors $\rho \in \Delta(\Delta(\Omega))$ has full support in the sense that $\Omega = \cup_{\pi \in \text{supp}(\rho)} \text{supp}(\pi)$, it defines by Bayes' rule a (conditional) prior over priors in every conditioning event; thus we can obtain $\rho_A = \text{BU}(\rho, A)$ for all $A \in \Sigma$. Therefore, we have a minimal CHTS representation $(u, \rho, \boldsymbol{\epsilon})$ where $\boldsymbol{\epsilon} = \{\epsilon_A\}_{A\in\Sigma}$, owing to the fact that conditional second-order beliefs must be consistent with the common prior ρ.

Next, suppose $\{\succeq_A\}_{A\in\Sigma}$ also satisfies CDC and is represented by a CHTS representation $(u, \rho, \boldsymbol{\epsilon})$, via an HTCPS $\{\pi_A\}_{A\in\Sigma}$. For any $A', A \in \Sigma$ with $A' \supseteq A$, CDC implies $\pi_A = \text{BU}(\pi_{A'}, A)$ if $\pi_{A'}(A) > 0$. By setting $\epsilon_{A'} = 0$ for all $A' \in \Sigma$, we have a minimal CHTS representation $(u, \rho, \mathbf{0})$. Now, suppose $(u, \rho, \mathbf{0})$ represents $\{\succeq_A\}_{A\in\Sigma}$. Consider arbitrary $A, A' \in \Sigma$ with $A \subseteq A'$ and A is not $\succeq_{A'}$-null. Then, $\pi_{A'}(A) > 0$. By the CHTS representation, $\pi_A = \text{BU}(\pi_{A'}, A)$. Therefore, $\{\succeq_A\}_{A\in\Sigma}$ satisfies CDC. $\square$

Proof of Corollary 1: "If": Suppose $\{\succeq_A\}_{A\in\Sigma}$ has a conditional-probability-system expected utility representation. Then, there exist a nonconstant affine function $u : X \to \mathbb{R}$ and a CPS $\{\pi_A\}_{A\in\Sigma}$ such that for any $f, g \in \mathcal{F}$

$$f \succeq_A g \Leftrightarrow \sum_{\omega\in\Omega} \pi_A(\omega)\, u(f(\omega)) \geq \sum_{\omega\in\Omega} \pi_A(\omega)\, u(g(\omega)).$$

Construct a prior over priors $\mathring{\rho} \in \Delta(\Delta(\Omega))$ as follows:

$$\mathring{\rho} = \frac{1-\delta}{(1-\delta^{K+1})} \sum_{k=0}^{K} \delta^k \pi^k$$

where $\delta > 0$ is sufficiently small, $\pi^0 = \pi_\Omega$, and for $k = 1, \cdots, K$,

$$\pi^k = \begin{cases} \pi_{\Omega \setminus \left[\cup_{\ell=0}^{k-1} \text{supp}\left(\pi^\ell\right)\right]} & \text{if } \Omega \neq \left[\cup_{\ell=0}^{k-1} \text{supp}\left(\pi^\ell\right)\right] \\ \pi^{k-1} & \text{if } \Omega = \left[\cup_{\ell=0}^{k-1} \text{supp}\left(\pi^\ell\right)\right] \end{cases}.$$

By construction, it is easily verified that $(u, \mathring{\rho}, \mathbf{0})$ is a CHTS representation for $\{\succeq_A\}_{A\in\Sigma}$. By Theorem 1, the class of preferences relations $\{\succeq_A\}_{A\in\Sigma}$ satisfies WbP, Consequentialism, and CDC.

"Only if": Let $\{\succeq_A\}_{A\in\Sigma}$ satisfy WbP, Consequentialism, and CDC. By Lemma 1(ii) and Theorem 1, the class of preferences relations $\{\succeq_A\}_{A\in\Sigma}$ admits a CHTS representation $(u, \rho, \mathbf{0})$. Thus, there exist a nonconstant affine function $u : X \to \mathbb{R}$ and an HTCPS $\{\pi_A\}_{A\in\Sigma}$ such that for any $f, g \in \mathcal{F}$

$$f \succeq_A g \Leftrightarrow \sum_{\omega\in\Omega} \pi_A(\omega)\, u(f(\omega)) \geq \sum_{\omega\in\Omega} \pi_A(\omega)\, u(g(\omega)).$$

Since $\boldsymbol{\epsilon}=\mathbf{0}$, for any $A, B, C \in \Sigma$, $\pi_A(C) = \frac{\pi_B(A\cap C)}{\pi_B(A)}$ if $A \subseteq B$ and $\pi_B(A) > 0$. Therefore, $\pi_B(C) = \pi_B(A)\,\pi_A(C)$, $\forall C \subseteq A \subseteq B$. That is, $\{\pi_A\}_{A\in\Sigma}$ is a CPS. $\square$

Proof of Corollary 2: The "If" part follows immediately from the second part (for $\boldsymbol{\epsilon}=\mathbf{0}$) of CHTS Representation Theorem.

"Only if": Let $\{\pi_A\}_{A\in\Sigma}$ be an HTCPS in a CHTS representation. Suppose $\{\pi_A\}_{A\in\Sigma}$ is a CPS on Ω. We adopt the construction of $\mathring{\rho}$ in the "If" part in the proof of Corollary 1. Clearly, $\mathring{\rho} \in \Delta(\Delta(\Omega))$ is a "partitional" prior over priors. By construction of $\mathring{\rho}$, for any $A \in \Sigma$, $\{\pi_A^*\} = \{\pi_A\} = \arg\max_{\pi\in\Delta(\Omega)} \mathring{\rho}(\pi)\,\pi(A)$. Therefore, $\pi_A =$BU(π_A^*, A), $\forall A \in \Sigma$. Since $\{\pi_A\}_{A\in\Sigma}$ is a CPS on Ω, for any $A, A' \in \Sigma$ satisfying $A \subseteq A'$, we have

$$\pi_{A'}(A\cap C) = \pi_{A'}(A)\,\pi_A(C), \forall C \subseteq \Omega;$$

that is, $\pi_A =$BU$(\pi_{A'}, A)$ if $\pi_{A'}(A) > 0$. $\square$

Proof of Corollary 3: By Ortoleva's (2012) Theorem 1, (1)$\Longleftrightarrow$(2). By the CHTS Representation Theorem, (2) $\Longleftrightarrow$(3). $\square$

To prove Theorem 3, we need the following lemma. Consider an HTM updating rule, $\boldsymbol{\varphi} = \{\varphi_A\}_{A\in\Sigma}$, for a belief-HTM $(\rho, \boldsymbol{\epsilon}, \boldsymbol{\pi})$.

Lemma 2. Suppose $A, \mathcal{A} \in \Sigma$ and $A \cap \mathcal{A} \neq \emptyset$. Then, $\varphi_A(\pi_A, \mathcal{A}) = \pi_{A\cap\mathcal{A}}$.

Proof of Lemma 2: Because $A, \mathcal{A} \in \Sigma$ and $A \cap \mathcal{A} \neq \emptyset$,

$$\begin{aligned}
\varphi_A(\pi_A, \mathcal{A}) &= \begin{cases} \mathrm{BU}(\pi_A, \mathcal{A}) & \text{if } \pi_A(\mathcal{A}) > \epsilon_A \\ \pi_{A \cap \mathcal{A}} & \text{if } \pi_A(\mathcal{A}) \leq \epsilon_A \end{cases} \\
&= \begin{cases} \mathrm{BU}(\pi_A, A \cap \mathcal{A}) & \text{if } \pi_A(A \cap \mathcal{A}) > \epsilon_A \\ \pi_{A \cap \mathcal{A}} & \text{if } \pi_A(A \cap \mathcal{A}) \leq \epsilon_A \end{cases}.
\end{aligned}$$

By CHTS Representation Theorem, $\pi_{A \cap \mathcal{A}} =BU(\pi_A, A \cap \mathcal{A})$ if $\pi_A(A \cap \mathcal{A}) > \epsilon_A$. Therefore, $\varphi_A(\pi_A, \mathcal{A}) = \pi_{A \cap \mathcal{A}}$. $\square$

Proof of Theorem 3: For all $A \in \Sigma$ and $B, C \in \Sigma_A$ with $B \cap C \neq \emptyset$, by Lemma 2,

$$\begin{cases} \varphi_B(\varphi_A(\pi_A, B), C) = \varphi_B(\pi_{A \cap B}, C) = \varphi_B(\pi_B, C) = \pi_{B \cap C} \\ \varphi_A(\pi_A, B \cap C) = \pi_{A \cap B \cap C} = \pi_{B \cap C} \\ \varphi_C(\varphi_A(\pi_A, C), B) = \varphi_C(\pi_{A \cap C}, B) = \varphi_C(\pi_C, B) = \pi_{C \cap B} \end{cases}.$$

Thus, for all $A \in \Sigma$ and $B, C \in \Sigma_A$ with $B \cap C \neq \emptyset$,

$$\varphi_B(\varphi_A(\pi_A, B), C) = \varphi_A(\pi_A, B \cap C) = \varphi_C(\varphi_A(\pi_A, C), B).$$

Now, assume that $\pi \in \Delta(\Omega)$ such that $\pi = \pi_\Omega$ in a belief-HTM $(\rho, \boldsymbol{\epsilon}, \boldsymbol{\pi})$. Letting $A = \Omega$, we obtain

$$\varphi_B(\varphi_\Omega(\pi_\Omega, B), C) = \pi_{B \cap C} = \varphi_C(\varphi_\Omega(\pi_\Omega, C), B),$$

for all $B, C \in \Sigma$ with $B \cap C \neq \emptyset$.$\square$

Proof of Corollary 4: Let $A \in \Sigma$ and $B, C \in \Sigma_A$. Since $\pi_A(B \cap C) > \epsilon_A$, $B \cap C \neq \emptyset$. By Theorem 3,

$$\varphi_B(\varphi_A(\pi_A, B), C) = \varphi_A(\pi_A, B \cap C) = \varphi_C(\varphi_A(\pi_A, C), B).$$

Again, since $\pi_A(B \cap C) > \epsilon_A$, $\varphi_A(\pi_A, B \cap C) =BU(\pi_A, B \cap C)$. Thus,

$$\varphi_B(\varphi_A(\pi_A, B), C) = \mathrm{BU}(\pi_A, B \cap C) = \varphi_C(\varphi_A(\pi_A, C), B),$$

for all $A \in \Sigma$ and $B, C \in \Sigma_A$ with $\pi_A(B \cap C) > \epsilon_A$. $\square$

References

[1] Anscombe, Frank J., and Robert J. Aumann. 1963. "A Definition of Subjective Probability." The Annals of Mathematical Statistics 34, 199–205.

[2] Aumann, Robert J. 1995. "Backward Induction and Common Knowledge of Rationality." Games and Economic Behavior 8, 6–19.

[3] Basu, Pathikrit. 2019. "Bayesian Updating Rules and AGM Belief Revision." Journal of Economic Theory 179, 455–475.

[4] Battigallia, Pierpaolo, and Marciano Siniscalchi. 2002. "Strong Belief and Forward Induction Reasoning." Journal of Economic Theory 106, 356–391.

[5] Benjamin, Daniel. 2019. "Errors in Probabilistic Reasoning and Judgment Biases," in: Handbook of Behavioral Economics-Foundations and Applications 2 (Eds. by B. D. Bernheim, S. DellaVigna, and D. Laibson). North-Holland: Elsevier, 69–186.

[6] Blume, Lawrence, Adam Brandenburger, and Eddie Dekel. 1991. "Lexicographic Probabilities and Choice under Uncertainty." Econometrica 59, 61–79.

[7] Börgers, Tilman. 1994. "Weak dominance and approximate common knowledge." Journal of Economic Theory 64, 265–276.

[8] Brandenburger, Adam, Amanda Friedenberg, and H. Jerome Keisler. 2008. "Admissibility in Games." Econometrica 76, 307–352.

[9] Dominiak, Adam, Matthew Kovach, and Gerelt Tserenjigmid. 2021. "Minimum Distance Belief Updating with General Information." Virginia Tech, Working paper.

[10] Epstein, Larry G. 2006. "An Axiomatic Model of Non-Bayesian Updating." Review of Economic Studies 73, 413–36.

[11] Epstein, Larry G., and Michel Le Breton. 1993. "Dynamically Consistent Beliefs Must Be Bayesian." Journal of Economic Theory 61, 1–22.

[12] Epstein, Larry G., and Martin Schneider. 2003. "Recursive Multiple-Priors." Journal of Economic Theory 113, 1–31.

[13] Epstein, Larry G., Jawwad Noor, and Alvaro Sandroni. 2008. "Non-Bayesian Updating: A Theoretical Framework." Theoretical Economics 3, 193–229.

[14] Galanis, Spyros. 2021. "Dynamic Consistency, Valuable Information and Subjective Beliefs." Economic Theory 71, 1467–1497.

[15] Ghirardato, Paolo. 2002. "Revisiting Savage in a Conditional World." Economic Theory 20, 83–92.

[16] Gilboa, Itzhak, and David Schmeidler. 1989. "Maxmin Expected Utility with Non-Unique Prior." Journal of Mathematical Economics 18, 141–53.

[17] Gilboa, Itzhak, and David Schmeidler. 1993. "Updating Ambiguous Beliefs." Journal of Economic Theory 59, 33–49.

[18] Halpern, Joseph. 2010. "Lexicographic Probability, Conditional Probability, and Nonstandard Probability." Games and Economic Behavior 68, 155–179.

[19] Hanany, Eran, and Peter Klibanoff. 2007. "Updating Preferences with Multiple Priors". Theoretical Economics, 3, 261–298.

[20] Karni, Edi, and Marie-Louise Vierø. 2013. "Reverse Bayesianism: A Choice-Based Theory of Growing Awareness." American Economic Review 103, 2790-2810.

[21] Karni, Edi, and Marie-Louise Vierø. 2017. "Awareness of unawareness: A theory of decision making in the face of ignorance." Journal of Economic Theory 168, 301–328.

[22] Ke, Shaowei, Brian Wu, and Chen Zhao. 2022. "Learning from a Black Box." University of Michigan, Working paper.

[23] Kreps, David M., and Robert Wilson. 1982: "Sequential Equilibria." Econometrica 50, 863–894.

[24] Myerson, Roger B. 1986a. "Axiomatic Foundations of Bayesian Decision Theory." Northwestern University, Center for Mathematical Studies in Economics and Management Science Discussion Paper 671.

[25] Myerson, Roger B. 1986b. "Multistage Games with Communication." Econometrica 54, 323–58.

[26] Ortoleva, Pietro. 2012. "Modeling the Change of Paradigm: Non-Bayesian Reactions to Unexpected News." American Economic Review 102, 2410–2436.

[27] Pires, Cesaltina Pacheco. 2002. "A Rule for Updating Ambiguous Beliefs." Theory and Decision 53 (2), 137–152.

[28] Sadler, Evan. 2021. "A Practical Guide to Updating Beliefs from Contradictory Evidence." Econometrica 89, 415–436.

[29] Samuelson, Larry. 1992. "Dominated Strategies and Common Knowledge." Games and Economic Behavior 4, 284–313.

[30] Savage, Leonard J. 1954. The Foundations of Statistics. New York: Wiley.

[31] Schipper, Burkhard C. 2022. "Predicting the Unpredictable under Subjective Expected Utility." UC_Davis, Working paper.

[32] Siniscalchi, Marciano. 2011. "Dynamic Choice under Ambiguity." Theoretical Economics 6, 379–421.

[33] von Neumann, John, and Oskar Morgenstern. 1944. Theory of Games and Economic Behavior. Princeton, NJ: Princeton University Press.

[34] Zhao, Chen. 2018. "Representativeness and Similarity." Working Paper, The University of Hong Kong.

[35] Zhao, Chen. 2022. "Pseudo-Bayesian Updating." Theoretical Economics 17, 253–289.