# Learning and Stochastic Dominance with Applications to Subjective Valuation of Options

Ehud Lehrer* Tao Wang†

September 3, 2023

**Abstract**

In the context of a binomial tree model, the decision maker has a subjective belief about the probability associated with upward and downward movements. Within this dynamic framework, we examine when various stochastic dominance relations among priors are preserved after conditioning on histories of observations, and when can we guarantee that the probability distributions over a certain set of histories induced by different priors satisfy first-order stochastic dominance relation. These stochastic dominance relations include first-order and likelihood-ratio dominance, and (reverse) hazard-rate dominance. As an application, we explore the consequences of diverse stochastic dominance relations on the subjective evaluation of both European and American options.

**Keywords:** Binomial tree model; Bayesian learning; Stochastic dominance; Option valuation.
**JEL Codes:** D80, D83.

*Department of Economics and Finance, Durham University, Mill Hill Lane, Durham DH1 3LB, UK. Email: ehud.m.lehrer@durham.ac.uk.

†International School of Economics and Management, Capital University of Economics and Business, Beijing, China. Email: tao.wang.nau@hotmail.com.

# 1 Introduction

In classical dynamic Bayesian decision problems involving uncertainty, a decision maker (referred to as DM) is faced with an unknown state of nature about which she holds an initial prior belief. As the decision scenario unfolds, the DM gathers information that assists in updating her belief about the state, thereby enabling her to maker better decisions. However, even when confronted with identical decision scenarios, DMs often possess distinct prior beliefs. These differences can be attributed to various factors, such as diverse past experiences and individual dispositions spanning from optimism to pessimism. Consequently, individuals may assign differing probabilities to future events, subsequently influencing their behavioral responses.

This paper aims to analyze the implications of different stochastic dominance relations among prior beliefs in Bayesian learning, focusing notably on the applications in terms of subjective valuation of options. In the Bayesian learning framework, a history of observations serves a dual role: it helps the DM to update her belief about the unknown state of nature, and it determines the DM's payoff. Drawing from the dual role histories play, we consider two fundamental facets of Bayesian learning: the stochastic dominance relation between the prior belief over a set of histories and between the posterior belief conditional on this set. While the probability distribution over histories is useful for evaluating the expected payoffs, the posterior belief is important for calculating continuation payoffs, and ultimately for the decision making. To address these questions, we investigate four most prominent stochastic dominance relations that compare the "magnitude" or "location" of random variables,[1] including first-order stochastic dominance, hazard-rate dominance, reverse hazard-rate dominance, and likelihood-ratio dominance.

As an application, we examine the connections between the various forms of stochastic dominance relations and the valuation of European and American options. The conventional no-arbitrage approach to option pricing operates on the premise that it is possible to construct a portfolio comprising traded financial assets that mirrors the price fluctuations of the underlying asset. Consequently, the subjective beliefs held by option holders regarding the unknown state play no role in option valuation. However, this method assumes a market rich enough in risky assets, which can be quite demanding.

By contrast, our focus lies in highlighting the significance of prior beliefs in a

[1]See, e.g., p. 3, Chapter 1 of Shaked & Shanthikumar (2007). We exclude second- and third-degree stochastic dominance, for which Levy (2006) offers a comprehensive exploration, coupled with applications in the context of investment decisions within uncertain environments. However, Levy's work does not contain an analysis of (reverse) hazard-rate dominance and likelihood-ratio dominance, both of which constitute central chapters of our paper.

decision-theoretic framework, and concentrate on the subjective valuation of options. This approach proves particularly relevant when considering the valuation of certain *real options*[2], wherein not every derivative can be replicated through portfolios. Our findings reveal that first-order stochastic dominance and (reverse) hazard-rate dominance exhibit a relationship with the valuation of European options, while likelihood-ratio dominance displays a connection with the valuation of American options.

Our analysis is conducted mainly within the framework of the binomial tree model, which boasts a broad spectrum of applications.[3] In this model, a stochastic outcome, either an "up" or "down" movement, is generated independently in each period, conditioned on the state. The probability $p$ of observing an "up" outcome remains unknown and constant over time. The value of $p$ can reflect various factors such as the profitability of an investment, the overall market conditions, and more. The DM holds a prior belief about the value of $p$, captured by a cumulative distribution function (CDF) on the interval $[0, 1]$. This CDF can be either continuous or discrete, and is allowed to have any number of atoms. Each observed sequence of outcomes, referred to as a history, progressively reveals information about the underlying state, enabling the DM to revise her belief, and thereby affecting her payoff.

Our initial set of findings revolves around first-order stochastic dominance. Consider two priors, denoted as $F$ and $G$, that do not put all probability on 0 or 1.[4] We show that $F$ first-order stochastically dominates $G$ if and only if for any period $t$, the probability distribution of the $t + 1$ distinct empirical frequencies[5] induced by $F$ first-order stochastically dominates that induced by $G$. Consequently, if the DM's payoff function is monotonically increasing in empirical frequencies, this equivalence implies that for any period $t$, a prior that exhibits first-order stochastic dominance would yield a greater expected payoff. Utilizing this equivalence, we demonstrate that $F$ first-order stochastically dominates $G$ if and only if the expected value of any European call option is higher under $F$ than under $G$.

We next consider two important stochastic dominance relations among prior beliefs: hazard-rate dominance and reverse hazard-rate dominance; both are stronger than first-

[2]Real options differ from conventional financial options in that they involve tangible assets such as land, buildings, inventory, etc. Some investment problems involving sunk costs could also be classified as real options. Consequently, real options are often not traded as securities. Furthermore, management cannot quantify uncertainty in terms of volatility and must instead rely on their perceptions of uncertainty.

[3]For instance, it serves as a foundational framework in diverse areas such as option pricing (e.g., Cox et al., 1979; Shreve, 2004) and Bayesian learning (e.g., Bikhchandani et al., 1992; Müller & Scarsini, 2002; Ifrach et al., 2019).

[4]This assumption ensures that every history is observed with a positive probability.

[5]These empirical frequencies denote the occurrences of "up" outcomes. In period $t$, there are $t+1$ empirical frequencies: $\frac{0}{t}, ..., \frac{t}{t}$.

order stochastic dominance. They capture different kinds of optimistic attitudes toward risk. We find that the hazard-rate dominance (reverse hazard-rate dominance) relation is preserved under Bayesian learning, conditional on any event whose probability is increasing (decreasing) in the state of nature. This property not only holds in the binomial tree model, but it extends also to the general Bayesian learning framework. The concept of hazard-rate dominance implies that upon observing a favorable event—referred to as "good news"—which is more likely to get realised under a more favorable state, such as a surge in sales, the DM maintains a sense of optimism concerning that state, as characterized by this concept. This property distinguishes it from first-order stochastic dominance, as the latter lacks this specific attribute after conditioning. Furthermore, consider a set of equal-length histories whose empirical frequencies are above a given threshold. It is shown that a hazard-rate dominating prior yields a first-order stochastically dominating probability distribution over this set of histories.

The importance of these results is that they bring forth a consequential insight regarding the valuation of European options: we show that $F$ hazard-rate dominates $G$ if and only if, for every European call option, the option value – under the condition that it is exercised – is greater under $F$ than under $G$.

The third set of findings centers around the concept of likelihood-ratio dominance, a form of stochastic dominance that is stronger than the other types explored in this study, and indeed yields stronger results. Mirroring the results established for (reverse) hazard-rate dominance, the likelihood-ratio dominance relation is preserved under conditioning on any event that occurs with positive probability. This result illuminates the connections that exist between likelihood-ratio dominance and the valuation of American options. Specifically, we establish that $F$ likelihood-ratio dominates $G$ if and only if for every American call option, the option's value associated with $F$ exceeds that associated with $G$, when conditioning on the event that the option remains unexercised. Moreover, in cases where $F$ likelihood-ratio dominates $G$, if it is optimal to exercise under $F$, the same holds true under $G$. This conveys that for any American call option, the set of histories leading to exercising of the option is smaller under the likelihood-ratio dominating prior. Hence holders of American options with likelihood-ratio dominating prior beliefs are more patient.

*Related Literature.* Our paper is related to the literature that studies stochastic dominance in the Bayesian learning framework. One strand of literature focuses on the stochastic dominance relations among posteriors induced by different samples, while holding the prior fixed. Whitt (1979) considers the influence of samples on posterior distribution of the entire population as well as the unsampled population in the binomial tree model similar to ours. Whitt's work lies within the domain of likelihood-ratio dominance. In the same vein, Fahmy et al. (1982) analyse the likelihood-ratio dominance of posteriors induced by histories of different empirical frequencies when the prior is fixed. Milgrom (1981) shows that a signal $x$ yields a first-order stochastically dominating posterior belief compared to another signal $y$ for every prior if and only if, the density function of the signal conditional on state satisfies the strict monotone likelihood-ratio property (see also Chambers & Healy (2009) for a related discussion). Charness et al. (2007) conduct an experimental study to test whether individuals and groups behave in accordance with first-order stochastic dominance and Bayesian updating when making decisions.

These studies delve into the impact of various histories while keeping the prior belief fixed. In contrast, our paper examines different priors and their influence on the stochastic dominance relationship of posterior distributions, as well as on the probability distributions over specific set of histories. Furthermore, while the previously mentioned papers exclusively focus on likelihood-ratio dominance and first-order stochastic dominance, our analysis expands to encompass a broader array of stochastic dominance relationships. This includes exploring hazard-rate dominance and reverse hazard-rate dominance.

Another strand of literature that studies stochastic dominance and Bayesian learning considers the implications of different prior beliefs. Bikhchandani et al. (1992) consider a finite set of simple lotteries over a finite set of outcomes. Priors (and posteriors) are compound lotteries. For two priors $F$ and $G$, they find necessary and sufficient conditions under which after any history of observations, the reduced simple lottery under $F$ first-order stochastically dominates (or "Bayes first-order stochastic dominates", in the terminology of Bikhchandani et al., 1992) that under $G$. As an application, Bikhchandani & Sharma (1996) examine the effect of Bayes' first-order stochastic dominance among priors in terms of an optimal search problem when the underlying distribution of price is unknown. In comparison, we study the stochastic dominance relations of posteriors conditional on a set of histories instead of a single history. In addition, we also examine probability distributions on a set of histories induced by different priors.

Our paper also relates to the literature on real option valuation. In contrast to financial options, the underlying assets for real options are often tangible. As described in Dixit & Pindyck (1994), generally there are two approaches to real option valuation, depending on whether the risk of the underlying asset can be spanned by existing assets or not. In case one can find traded assets that exactly replicate the stochastic component of the underlying asset, the classical arbitrage-free option pricing approach can be applied (see, e.g., Cox et al., 1979; Shreve, 2004). Otherwise, we refer to the dynamic programming approach: assume the price of the underlying asset evolves according to a *known* stochastic process. The dynamic programming method is used to obtain the option value in a wide range or dynamic decision problems with irreversible actions under uncertainty, such as investment problems (see, e.g., Cukierman, 1980; Bernanke, 1983; Pindyck, 1991; Demers, 1991; Chetty, 2007). Our paper pertains to this second approach. In contrast to the literature, we consider the situation in which there is uncertainty in the stochastic process (i.e., the probability of moving up each period is unknown, instead of a known constant) and focus on different stochastic dominance relations among prior beliefs and investigate the implications for the option value. Stochastic dominance has been applied to option valuation when market completeness or frictionless trading assumptions are violated, for instance, due to the presence of rare events and proportional transaction costs. The idea is to obtain upper and lower bounds on admissible option prices (instead of a unique option price) that exclude *stochastically dominant strategies* (strategies that increase the utility of all traders). For option prices outside the bounds, every trader can have an arbitrage opportunity. Perrakis (2019) contains a comprehensive treatment of studies along this vein, with a focus on the second-degree stochastic dominance. By contrast, we focus on likelihood-ratio dominance and (reverse) hazard-rate dominance relations among prior and posterior beliefs about an unknown state of nature that governs the evolution of the market price.

The rest of the paper is organized as follows. Section 2 presents the binomial tree model. The following three sections are devoted to the discussion of first-order stochastic dominance, (reverse) hazard-rate dominance, and likelihood-ratio dominance relations among prior beliefs, respectively. The proofs and some supplementary results can be found in the Appendix.

# 2 The Binomial Tree Model

Consider a risk-neutral decision maker (DM) who faces an unknown state of nature $p \in [0, 1]$. The DM's prior belief about $p$ is captured by a CDF $F(p)$. Depending on the true state of nature, a random outcome (either "$U$" or "$D$") is observed in each period. Suppose that $\mathbb{P}(U) = p$ and outcomes observed in different periods are independent conditional on the state.

As an example, the symbols "$U$" and "$D$" can be understood as representing upward and downward shifts in a company's stock price, similar to the interpretation in the work by Cox et al. (1979). The parameter $p$ could signify the undisclosed performance of the company or the prevailing market conditions. In different scenarios, these symbols could also be associated with distinct meanings. For instance, they might represent changes in a company's sales—either an increase ("$U$") or a decrease ("$D$"). Alternatively, they could denote positive ("$U$") and negative ("$D$") consumer ratings, as seen in the research conducted by Ifrach et al. (2019)

A history of length $t$ is an ordered sequence of outcomes, and is denoted by $h_t$. Let $\mathscr{H}_t$ be the set of all histories of length $t$, and denote a subset of $\mathscr{H}_t$ by $H_t$. When a history $h_t$ contains $k$ observations of the outcome "$U$", we say that the *empirical frequency* (or simply *frequency*) of outcome "$U$" in $h_t$ is $\theta(h_t) := \frac{k}{t}$. Histories have two roles: when the dynamics unfolds, they reveal information to the DM about the unknown state $p$; moreover, they determine the payoff of the DM.

One of the most notable applications of the binomial tree model is in option pricing. As demonstrated by Cox et al. (1979), the celebrated Black-Scholes formula can be derived as an asymptotic case of the binomial tree model. However, the conventional theory of financial option pricing is centered around the fundamental concept of the *no-arbitrage principle*. This principle assumes the free tradability of the underlying asset in the market, positing that it is possible to construct a portfolio that perfectly mirrors the price fluctuations of the underlying asset. This methodology relies, firstly, on the availability of a suitably diverse array of markets dealing with risky assets – a requirement that can be quite demanding. Secondly, it hinges on the assumption that risk-free arbitrage opportunities do not exist. This latter assumption effectively nullifies the impact of the option holder's personal beliefs about the parameter $p$ on the determination of the option price. In contrast, to underscore the influence of prior beliefs, we embrace a decision-theoretic perspective. This approach proves particularly relevant in scenarios like the valuation of numerous real options, including various investment-decision problems involving sunk costs.

# 3 First-Order Stochastic Dominance and European Options

A key driver behind the widespread use of first-order stochastic dominance (FOSD) lies in the following well-known equivalent condition, articulated through expectations: $F \succeq_{\text{FOSD}} G$ if and only if, $\mathbb{E}_F(u) \geq \mathbb{E}_G(u)$, where $u$ is any increasing[6] function. This section of the study delves into the implications of the first-order stochastic dominance relation among prior distributions within the binomial tree model. We establish an equivalent condition for the relationship $F \succeq_{\text{FOSD}} G$, tailored to this specific model. Subsequently, we extend this understanding to the subject of European option valuation and thereby illustrate additional applications of this concept.

Notice that for a given period $t$, the prior belief $F$ induces a probability distribution

[6] In this context, "increasing" (or "decreasing") is used to indicate weakly increasing (or decreasing) functions.

over the $t+1$ empirical frequencies $\frac{0}{t}, ..., \frac{t}{t}$ via

$$\mathbb{P}_F(k,t) := \int_0^1 \binom{t}{k} p^k (1-p)^{t-k} \mathrm{d}F, \quad 0 \leq k \leq t. \tag{1}$$

In other words, $\mathbb{P}_F(k,t)$ is the probability of observing a history with empirical frequency $\frac{k}{t}$ in period $t$ under the prior belief $F$.

The following proposition gives an equivalent condition for $F \succeq_{\text{FOSD}} G$ expressed in terms of the stochastic dominance relation of probability distributions over empirical frequencies induced by different priors. The proof, as well as all others, is deferred to the Appendix.

**Proposition 1.** *Consider two different prior beliefs $F$ and $G$ about the state of nature $p$. Then $F \succeq_{\text{FOSD}} G$ if and only if for any period $t$, the probability distribution over the $t+1$ empirical frequencies under prior $F$ first-order stochastically dominates the probability distribution under $G$, or $\{\mathbb{P}_F(k,t)\}_{k=0}^t \succeq_{\text{FOSD}} \{\mathbb{P}_G(k,t)\}_{k=0}^t$.*

It is important to note that the first-order stochastic dominance relation is not preserved under conditioning, whether that involves conditioning on a subset of the support, on a single history, or on a set of histories. This fact is well-documented in the literature, as exemplified by studies like Bikhchandani et al. (1992). In contrast, stronger stochastic dominance relations—such as hazard-rate dominance, reverse hazard-rate dominance, and likelihood-ratio dominance—do exhibit this property. To further understand this issue, the reader is referred to Propositions 3 and 6 below.

**Application to Option and Forward Contract Valuation.** Proposition 1 naturally finds practical relevance in the valuation of European options and forward contracts. Consider a European call/put option, where owning a unit grants the holder the choice (but not the obligation) to buy/sell a specified asset unit at a predetermined strike price at a future time (known as the expiration date). Importantly, this choice is irrespective of the market price of the asset at that point. It is important to note that exercising this option is exclusively permissible on the expiration date, setting it apart from American call options that can be exercised at any point before expiration. Similarly, a forward contract determines both a termination date and a pre-established strike price. Nevertheless, upon reaching the termination date, the contract holder has no choice but to exercise at the pre-established strike price.

Consider a scenario where the market price evolves according to the binomial tree. A key assumption underlying this setup is that the price is contingent on the empirical frequency. More precisely, in period $t$, with $k$ instances of the outcome labeled as "$U$", the corresponding price is $S(k,t)$. Consequently, two equal-length histories with the same empirical frequency would result in identical prices. The function $S(k,t)$ is assumed to be monotonically increasing in $k$ for any fixed $t$. The conventional binomial option pricing model, such as outlined in Cox et al. (1979), assumes that the price moves up or down by a fixed fraction every period. Instead, in our setup the price function $S(k, t)$ is more flexible, accommodating for a broader set of options.

A unit of *European option* or *forward contract* is formally described by a triple $(T, \bar{S}, S(k, T))$, where $T$ denotes the termination date and $\bar{S}$ represents the predetermined strike price. Notably, $S(k, T)$ stands for the price at period $T$ given $k$ instances of the outcome "$U$". The probability of observing the outcome "$U$" is dictated by the unknown parameter $p$.

The ex-post value of the European call and put options at $(k, T)$ are, respectively, $v(k, T, S(k, T)) = \max\{0, S(k, T) - \bar{S}\}$ and $v(k, T, S(k, T)) = \max\{0, \bar{S} - S(k, T)\}$. For a forward contract, the ex-post value at $(k, T)$ is $v(k, T, S(k, T)) = S(k, T) - \bar{S}$. It follows from the monotonicity of $S(k, T)$ that $v(k, T)$ is increasing (decreasing) in $k$ for European call (put) options.

Suppose, in addition, that future payoffs are discounted by $\delta \in (0, 1]$. The ex-ante (subjective) value of the option when the option holder has prior belief $F$ is[7]

$$\begin{aligned} V_F(T, \bar{S}, S(k, T)) &:= \delta^T \int_0^1 \sum_{k=0}^{T} \binom{T}{k} p^k (1-p)^{T-k} v(k, T, S(k, T)) \mathrm{d}F(p) \\ &= \delta^T \sum_{k=0}^{T} \mathbb{P}_F(k, T) v(k, T, S(k, T)). \end{aligned} \tag{2}$$

The main finding of this section concerns the consequences of first-order stochastic dominance for the valuation of European call options and forward contracts.

**Theorem 1.** *Consider two beliefs $F$ and $G$ about the state of nature $p$. Then,*
*(i) $F \succeq_{\text{FOSD}} G$ if and only if the* ex-ante *value of every European call option is greater under $F$ than under $G$; and*
*(ii) $F \succeq_{\text{FOSD}} G$ if and only if the* ex-ante *value of every forward contract is greater under $F$ than under $G$.*

In the classical binomial option pricing theory (see, e.g., Cox et al., 1979), the price of the underlying asset at a node $(k, t)$ is given by $S(k, t) = S_0 \lambda^k \lambda^{-(t-k)} = S_0 \lambda^{2k-t}$, where $S_0$ is the initial price, and $\lambda > 1$ is a constant. In words, the price exhibits an increase or decrease by a fixed proportion with each iteration. Fix $S_0$, a price function can be characterized by $\lambda$. Consider a termination date $t$. Lemmas 3 and 4 in Appendix A.2 show that in order to determine whether $\{\mathbb{P}_F(k, t)\}_{k=0}^{t} \succeq_{\text{FOSD}} \{\mathbb{P}_G(k, t)\}_{k=0}^{t}$ holds or not, it suffices to consider any $t$ such price functions with different $\lambda$'s.

[7] Recall from Eq. (1) that $\mathbb{P}_F(k, T)$ is the probability of observing histories with frequency $\frac{k}{T}$ in period $T$ under prior belief $F$.

## 4 (Reverse) Hazard-Rate Dominance

The objective of this section is to investigate two additional stochastic orders among prior beliefs and their implications in the binomial tree model. We refer here to the hazard-rate and the reverse hazard-rate dominance orders (HRD and RHRD). It is well known that HRD and RHRD are stronger than FOSD. We show that under (R)HRD, one can obtain a stronger version of Proposition 1 and Theorem 1.

Consider two prior beliefs $F$ and $G$ on $[0, 1]$, and let $f$ and $g$ be their densities (Radon-Nikodym derivatives w.r.t. some common measure, not necessarily the Lebesgue measure). It includes cases in which $F$ and $G$ have atoms. To rule out uninteresting cases, we assume that $F$ and $G$ are *non-trivial*, namely, they are not degenerate at a single point, and the probability that $p \in (0, 1)$ (i.e., not 0 or 1) is positive. We know that $F$ *hazard-rate dominates* $G$ ($F \succeq_{\text{HR}} G$) if $\frac{f(p)}{1-F(p)} \leq \frac{g(p)}{1-G(p)}$. Similarly, $F$ *reverse hazard-rate dominates* $G$ ($F \succeq_{\text{RHR}} G$) if $\frac{f(p)}{F(p)} \geq \frac{g(p)}{G(p)}$. The following proposition presents two useful equivalent conditions for (R)HRD.

**Proposition 2.** *The following are equivalent conditions for $F \succeq_{\text{HR}} G$ (resp., $F \succeq_{\text{RHR}} G$):*

1. *For all functions $u$ and $w$ such that $w$ is non-negative and increasing, $u$ is increasing, and that the expectations under $F$ and $G$ exist,*

$$\mathbb{E}_F(uw)\mathbb{E}_G(w) \geq \mathbb{E}_G(uw)\mathbb{E}_F(w). \tag{3}$$

   *For* RHRD, *we only need to change $w$ to be non-negative and decreasing.*

2. *For any $p_H > p_L$, $[1 - F(p_H)][1 - G(p_L)] \geq [1 - G(p_H)][1 - F(p_L)]$ (resp., $F(p_H)G(p_L) \geq G(p_H)F(p_L)$ ).*

The first condition is due to Capéraà (1988). It is an equivalent condition for (R)HRD expressed in terms of expectations. Another way to express the first equivalent condition is as follows. Take any $w$ that is non-negative and increasing (resp., decreasing) such that $\mathbb{E}_F(w)$, $\mathbb{E}_G(w) > 0$. Define two new CDFs, $\hat{F}(p) := \frac{\mathbb{E}_F(w\mathbf{1}_{(-\infty,p]})}{\mathbb{E}_F(w)}$ and $\hat{G}(p) := \frac{\mathbb{E}_G(w\mathbf{1}_{(-\infty,p]})}{\mathbb{E}_G(w)}$. Then for any increasing and measurable function $u$, $\mathbb{E}_{\hat{F}}(u) \geq \mathbb{E}_{\hat{G}}(u)$, or $\hat{F} \succeq_{\text{FOSD}} \hat{G}$. It means that first-order stochastic dominance is preserved under a certain change of measure.

In the environment of Bayesian learning, since the posterior beliefs and conditional expectations can be expressed in the form $\frac{\mathbb{E}_F(uw)}{\mathbb{E}_F(w)}$, where $u$ and $w$ are functions with certain properties, this characterization also has natural applications (see, e.g., Proposition 4).

In comparison, the second equivalent condition, which is well-known (see, e.g., Shaked & Shanthikumar (2007)), is phrased in terms of CDFs. It has a straightforward geometric implication. To see it, let us consider the transformation (correspondence) $\varphi : [0,1] \rightrightarrows [0,1]$ such that $F(p) = \varphi(G(p))$.[8]

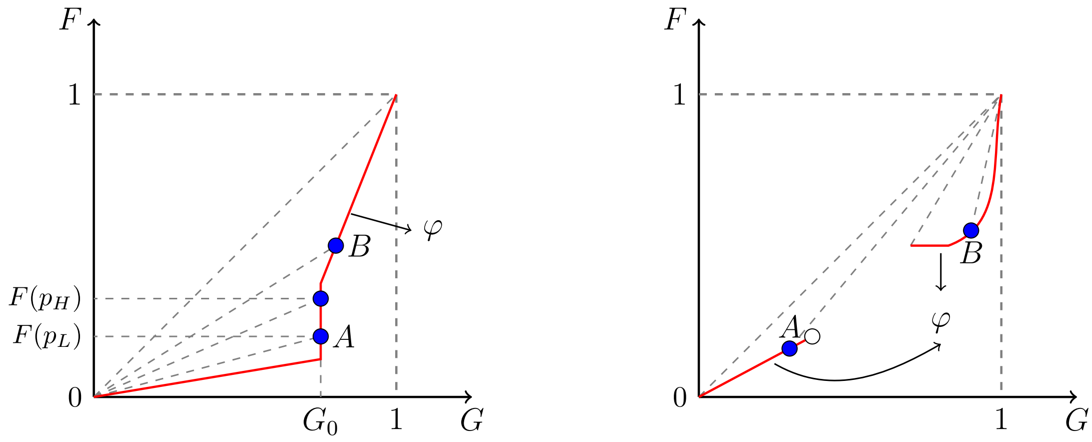


Figure 1: A graphical illustration of RHR and HR dominance.

Take any point on the graph of $\varphi$ and connect it to the origin by a segment. Condition 2 of Proposition 2 implies that if $F \succeq_{\text{RHR}} G$, then the slope of the line segment increases as one moves the point along the graph of $\varphi$ toward $(1,1)$ (see Panel (a) of Figure 1). In other words, imagine we put a light bulb at the origin, then the light can reach every point on $\varphi$ from above the graph of $\varphi$. The case when $F \succeq_{\text{HR}} G$ is similar. It implies that the slope of the line segment connecting a point on the graph of $\varphi$ and $(1,1)$ is increasing as one moves the point toward $(1,1)$.

Using Proposition 2, one can establish the following proposition, which is a main result of this section. It presents a set of conditions equivalent to $F \succeq_{\text{HR}} G$ (resp., $F \succeq_{\text{RHR}} G$) in the setup of Bayesian learning framework.

**Proposition 3.** *Consider two non-trivial priors $F$ and $G$ on $[0,1]$. The following conditions are equivalent:*

[8]Indeed, define $\varphi(y) = F(G^{-1}(y))$, where $y$ is in the range of $G$ (since $G$ may have atoms, the range of $G$ may be a proper subset of $[0,1]$) and $G^{-1}(y) = \{x \in [0,1] \mid G(x) = y\}$. The reason why $\varphi(\cdot)$ is a correspondence rather than a function is that $G$ might be constant on some interval, while $F$ is strictly increasing on the same interval. Such a case is illustrated in Panel (a) of Figure 1. In the figure, $G$ equals $G_0$ on $[p_L, p_H]$, while $F(p_H) > F(p_L)$, and $\varphi$ maps $G_0$ to a closed interval containing $F(p_L)$ and $F(p_H)$. Also, note that $F$ or $G$ may have atoms (hence the densities w.r.t. the Lebesgue measure may not exit) but nevertheless, $F \succeq_{\text{RHR}} G$ or $F \succeq_{\text{HR}} G$ can still hold (see Panel (b) of Figure 1).

1. *$F \succeq_{\mathrm{HR}} G$;*

2. *For a set of histories $H$ such that the probability of observing $H$ is increasing in $p$, the posterior beliefs conditional on $H$ satisfy $F|H \succeq_{\mathrm{HR}} G|H$.*

3. *For any period $t$ and any level of empirical frequency $\theta \in [0,1]$, consider the set of histories $H_t^+(\theta) := \{h_t \in \mathscr{H}_t | \theta(h_t) \geq \theta\}$. The posteriors conditional on $H_t^+(\theta)$ satisfy $F|H_t^+(\theta) \succeq_{\mathrm{FOSD}} G|H_t^+(\theta)$.*

4. *Consider the same set of length-$t$ histories $H_t^+(\theta)$ as in 3. The probability distribution over $H_t^+(\theta)$ under $F$ FOSD the probability distribution over $H_t^+(\theta)$ under $G$.*

*For $F \succeq_{\mathrm{RHR}} G$, we replace the set of histories $H$ in 2 with a set of histories whose probability is decreasing in $p$, and replace the set of histories $H_t^+(\theta)$ in 3 and 4 with $H_t^-(\theta) := \{h_t \in \mathscr{H}_t | \theta(h_t) \leq \theta\}$.*

The significance of Proposition 3 is that it establishes the equivalence of several conceptually distinct conditions: Condition 1 is concerned with stochastic dominance relations among prior beliefs. Conditions 2 and 3 focus on the stochastic dominance relations among posterior beliefs. In particular, condition 2 implies that the HRD (or RHRD) relation among prior beliefs is preserved conditional on certain histories. The third condition, which is weaker than the second one, is sufficient to ensure that the (R)HRD relation holds among prior beliefs. Condition 4 compares probability distributions over a set of histories induced by different priors.

In the second condition, the set of histories $H$ under consideration include many natural cases. In particular, by Lemma 1 in the Appendix, the probability of observing $H_t^+(\theta)$ in the third condition is increasing in $p$. The equivalence between the first two conditions has natural applications in Bayesian learning. We can extend the equivalence between the first two conditions to the general Bayesian learning framework by using Proposition 2.

**Proposition 4.** *Let $p \in \mathbb{R}$ be an unknown state of nature, and let $F(p)$, $G(p)$ be two prior beliefs. Then $F(p) \succeq_{\mathrm{HR}} G(p)$ (resp., $F(p) \succeq_{\mathrm{RHR}} G(p)$) if and only if, for any Blackwell experiment $\mathscr{E} = (\pi(s|p))_{s \in S}$,[9] and for any event $E \subseteq S$ whose probability is positive under $F$ and $G$ and is increasing (resp., decreasing) in $p$, the posterior beliefs conditional on $E$ satisfy $[F|E](p) \succeq_{\mathrm{HR}} [G|E](p)$ (resp., $[F|E](p) \succeq_{\mathrm{RHR}} [G|E](p)$).*

[9] Here $\pi(s|p)$ is the probability distribution over signals conditional on $p$. For a more comprehensive discussion of Blackwell experiment, see Blackwell (1953).

It is worth noting that, generally speaking, the first-order stochastic dominance relation is not preserved under conditioning, as is well-known in the literature (see, e.g., Bikhchandani et al., 1992). This is by contrast to stronger stochastic dominance relations, such as hazard-rate dominance and reverse hazard-rate dominance relations. Proposition 4 characterizes a class of events conditional on which the posteriors preserve the (reverse) hazard-rate dominance relation among priors.

Finally, condition 4 allows us to compare the expected payoffs under different prior beliefs. In particular, for a fixed $t$, if the payoff function is increasing in empirical frequency, then condition 4 implies that from the *ex-ante* perspective[10], a hazard-rate dominating prior yields a greater expected payoff over the set of histories $H_t^+(\theta)$.

Frequently, there is an interest in scrutinizing the posterior beliefs concerning a collection of histories, given that this precise collection has occurred. To motivate this notion, consider a scenario where, at time $t$, the DM is told that the realized history is in the set $H_t$, but she does not know which particular history has been realized. To draw a parallel, imagine an investor who lacks precise knowledge of an asset's value but possesses the information that the asset's value surpasses a certain threshold. The information that the actual history belongs to the set $H_t$ reveals additional information about the unknown state and allows the DM to update her belief from $F$ to $F|H_t$ about specific histories within $H_t$.

The term *ex-post* is assigned to the probability distribution over the set of histories $H_t$ which is assessed through the revised belief $F|H_t$. It differs from the ex-ante probability distribution over $H_t$ in that the ex-post uses the extra information embedded in that $H_t$ actually occurred. The ex-ante probability distribution, in contrast, is calculated using the prior belief $F$ alone, and lacks the additional information that $H_t$ has been realized.

**Corollary 1.** *Suppose that two prior beliefs satisfy $F \succeq_{\mathrm{HR}} G$. Then, given the set of histories $H_t^+(\theta)$, the* ex-post *probability distribution over $H_t^+(\theta)$ under $F$* FOSD *the* ex-post *probability distribution $H_t^+(\theta)$ under $G$. In case $F \succeq_{\mathrm{RHR}} G$, the same statement holds when $H_t^-(\theta)$ replacing $H_t^+(\theta)$.*

The corollary is an immediate consequence of Proposition 3. By Lemma 1 and the equivalence of conditions 1 and 2, the posteriors satisfy $F|H_t^+(\theta) \succeq_{\mathrm{HR}} G|H_t^+(\theta)$. Now regard $F|H_t^+(\theta)$ and $G|H_t^+(\theta)$ as the new "priors", the equivalence of conditions 1 and 4 establishes the desired result.

**Application to Valuation of European Options.** We return to the European options discussed by the end of Section 3. Since both hazard-rate dominance and

[10] Note that in condition 4 of Proposition 3, the probability distributions over the set of histories $H_t^+(\theta)$ are evaluated using the prior beliefs.

reverse hazard-rate dominance are stronger than first-order stochastic dominance, both $F \succeq_{\mathrm{RHR}} G$ and $F \succeq_{\mathrm{HR}} G$ imply that the *ex-ante* expected option value of every European call option is greater under $F$ than under $G$. However, we can obtain a stronger result for (reverse) hazard-rate dominance.

**Theorem 2.** *Let $F$ and $G$ be two non-trivial prior beliefs on $[0,1]$. Then $F \succeq_{\mathrm{HR}} G$ if and only if for every European* call *option, the option value* conditional on the option is exercised *is greater under $F$ than under $G$. Similarly, $G \succeq_{\mathrm{RHR}} F$ if and only if for any European* put *option, the option value* conditional on the option is exercised *is greater under $F$ than under $G$.*

An European call (put) option is exercised at the expiration date when the price of the underlying asset exceeds (falls below) the strike price. In the binomial tree model, the histories that lead to exercise of the option are exactly those whose empirical frequencies are above (below) a threshold. Therefore, we can use the equivalence between the first and the fourth conditions of Proposition 2 to establish Theorem 2.

# 5 Likelihood-Ratio Dominance and American Options

Another stochastic dominance relation of cardinal importance in economics, management science and finance is the likelihood-ratio dominance (LRD). It is stronger than HRD and RHRD. In this section, we focus on the implication of LRD in Bayesian updating, with applications to valuation of American options. We know that for two probability distributions $F$ and $G$ (CDFs), $F$ likelihood-ratio dominates $G$, and then one writes $F \succeq_{\mathrm{LR}} G$, if their densities $f$ and $g$ (w.r.t. some dominating measure, not necessarily the Lebesgue measure) satisfy $f(p_H)g(p_L) \geq f(p_L)g(p_H)$, for any $p_H > p_L$. Analogous to the equivalent conditions for (R)HRD (Proposition 2), we can establish the following counterpart result for LRD.

**Proposition 5.** *The following are equivalent conditions for $F \succeq_{\mathrm{LR}} G$:*

1. *For any* increasing *function $u$ and any* non-negative *function $w$, with respect to which the expectations exist, it holds that*

$$\mathbb{E}_F(uw)\mathbb{E}_G(w) \geq \mathbb{E}_G(uw)\mathbb{E}_F(w). \tag{4}$$

2. *For any $p_1 > p_2 > p_3$,*

$$[F(p_1) - F(p_2)][G(p_1) - G(p_3)] \geq [F(p_1) - F(p_3)][G(p_1) - G(p_2)]. \tag{5}$$

The first equivalent condition is due to Lehrer & Wang (2023). In contrast to Proposition 2, for LRD, there is no monotonicity requirement for the "weight function" $w$. The second equivalent condition is well known, it can be found in Shaked & Shanthikumar (2007). It implies that $F$ can be obtained from $G$ via a convex transformation $\varphi$ (i.e., there is a convex function $\varphi : [0,1] \to [0,1]$ such that $F = \varphi(G)$).

The equivalent conditions of Proposition 5 are particularly useful when one studies conditioning and Bayesian updating. Using these conditions, one can establish the following set of equivalent conditions for LRD in the binomial tree model.

**Proposition 6.** *Let $F$ and $G$ be two non-trivial prior beliefs on $[0,1]$. The following conditions are equivalent:*

1. $F \succeq_{\mathrm{LR}} G$;
2. *for any set of histories $H$ (not necessarily of the same length), $F|H \succeq_{\mathrm{LR}} G|H$;*
3. *for any set of histories $H$, $F|H \succeq_{\mathrm{FOSD}} G|H$;*
4. *for any $t$ and any set of length-$t$ histories $H_t \subseteq \mathscr{H}_t$, the probability distribution over $H_t$ under $F$* FOSD *that under $G$.*

This proposition is the counterpart of Proposition 3 for (R)HRD. Compared to Proposition 3, the second and the third conditions of Proposition 6 impose no restriction on the set of histories under consideration. As the following proposition demonstrates, the equivalence between the first and the second conditions is not restricted to the binomial tree model: it extends to more general Bayesian learning setups.

**Proposition 7.** *Let $p \in \mathbb{R}$ be an unknown state of nature, and let $F(p)$, $G(p)$ be two prior beliefs. Then $F(p) \succeq_{\mathrm{LR}} G(p)$ if and only if, for any Blackwell experiment $\mathscr{E} = (\pi(s|p))_{s\in S}$, and any event $E \subseteq S$ that occurs with positive probabilities under $F$ and $G$, the posterior beliefs conditional on $E$ satisfy $[F|E](p) \succeq_{\mathrm{LR}} [G|E](p)$.*

The property of likelihood-ratio dominance (LRD) relation being preserved under conditioning holds significant importance within Bayesian learning context. The next discussion on American options exemplifies its practical application.

Analogous to Corollary 2, we may examine the ex-post probability distributions over a set of histories. By utilizing Proposition 5, we can obtain the following result.

**Corollary 2.** *Suppose that two prior beliefs satisfy $F \succeq_{\mathrm{LR}} G$. Then, given any set of histories $H_t$, the* ex-post *probability distribution over $H_t$ under $F$* FOSD *the* ex-post *probability distribution $H_t$ under $G$.*

**Application to American Option Valuation**. An American call/put option entitles its holder to the right, though not the obligation, to purchase/sell a unit of a specified asset at a predetermined strike price *on or before* a designated future time, referred to as the expiration date. This is true regardless of the market price of the asset at the moment of exercise. This grants American options a higher degree of flexibility compared to European options.

In this context, a history of observations serves a dual role. Firstly, it determines the payoff of the option if exercised. Secondly, it provides information about the unknown state $p$, enables the option holder to revise her belief and to determine the optimal exercise timing. Essentially, the option holder is facing an optimal stopping problem, where an optimal exercise strategy corresponds to an optimal stopping rule. Our primary goal is to explore the ramifications of the likelihood-ratio dominance (LRD) relation among prior beliefs in terms of the value of American options and the corresponding optimal exercising strategy.

Consider an American option characterized by $(T, \bar{S}, S(k,t))$, where $S(k,t)$ is increasing in $k$ for a given $t$. Let $V_F^T(k,t)$ be the value function (option value) given a history with frequency $\frac{k}{t}$ under prior belief $F$. The value function can be written recursively as

$$V_F^T(k,t) = \max\left\{S(k,t) - \bar{S}, \delta\left[\mathbb{P}_F(U|k,t)V_F^T(k+1,t+1) + \mathbb{P}_F(D|k,t)V_F^T(k,t+1)\right]\right\}, \quad (6)$$

where $\mathbb{P}_F(U|k,t)$ and $\mathbb{P}_F(D|k,t)$ are the probabilities of observing "$U$" and "$D$" in period $t+1$, respectively, conditional on a history with frequency $\frac{k}{t}$, and $\delta \in (0,1]$ is the discount factor.

We examine first how LRD relation among prior beliefs affects the optimal exercise strategy. According to the value function Eq. (6), the optimal exercise strategy is governed by the comparison of the expected stopping payoff and the discounted continuation value. It follows immediately from Proposition 6 that conditional on any history $h_t$, a LRD prior yields a greater ex-post option value. Thus, we can obtain the following result:

**Proposition 8.** *Let $F$ and $G$ be two non-trivial prior beliefs that satisfy $F \succeq_{\text{LR}} G$. Then, for every American call option, whenever it is optimal to exercise under $F$, it is optimal to exercise under $G$.*

Another way to phrase Proposition 8 is to say that whenever it is optimal to wait under $G$, it is optimal to wait under $F$. This implies that when faced with the same American option, the option holder with a LRD prior tends to be more patient.

Next, we consider the implication of a LRD prior belief in terms of option value. In particular, we compare the option values conditional on a set of histories. We say that a history $h_t$ is *alive*, if the option is not exercised when $h_t$ is observed. Similarly, we say that a set of length-$t$ histories $H_t$ is *alive*, if each history in $H_t$ is alive. Using Proposition 6, we can establish the following result.

**Theorem 3.** *Let $F$ and $G$ be two non-trivial prior beliefs. Then $F \succeq_{\mathrm{LR}} G$ if and only if for every American call option and every set of histories $H_t$ that is alive under both $F$ and $G$, the expected option value conditional on $H_t$ is greater under $F$ than under $G$.*

The sets of histories under consideration are rich enough to cover many interesting cases. For instance, the ex-ante value (i.e., conditional on the null history) and the value conditional on every history that is alive is greater under a LRD prior. For American put options, $F \succeq_{\mathrm{LR}} G$ implies that the option value is lower under $F$.

# 6 Conclusion

We examined the implications of several well-known stochastic dominance relations that compare the magnitude or location of random variables in the Bayesian learning framework, and investigated applications in terms of subjective valuation of European and American options. These stochastic dominance relations include likelihood-ratio dominance, (reverse) hazard-rate dominance, and first-order stochastic dominance. In particular, we discussed when the stochastic dominance relation among prior beliefs is preserved under conditioning, and when it can be guaranteed that the probability distributions over a certain set of histories induced by different prior beliefs satisfy first-order stochastic dominance relation.

We have found that if a prior $F$ hazard-rate dominates (resp., reverse hazard-rate dominates) another one $G$, then conditional on any event whose probability is increasing (resp., decreasing) with the state of nature, the posteriors preserve the hazard-rate dominance (resp., reverse hazard-rate dominance) relation. By contrast, the likelihood-ratio dominance is preserved conditional on any event that occurs with positive probability. These result holds not only in binomial tree models, but also in the general Bayesian learning framework.

In the binomial tree model, we demonstrated that first-order stochastic dominance and (reverse) hazard-rate dominance are closely related to the subjective valuation of European options, while likelihood-ratio dominance is related to the valuation of American options. Specifically, a prior belief $F$ dominates another one $G$ in the sense of first-order stochastic dominance if and only if the value of every European call option is greater under $F$ than under $G$. The belief $F$ hazard-rate dominates $G$ if and only if for every European call option, conditional on the option is exercised, $F$ yields greater expected exercise payoff. For likelihood-ratio, a prior likelihood-ratio dominates another one if and only if the former yields greater valuation for every American call option.

lt is worth mentioning that to emphasize the role of the prior beliefs, we focused on subjective valuation of options. This approach may be applicable to the valuation of some real options, such as some investment problems involving sunk costs. This is in contrast to the no-arbitrage approach adopted for the valuation of financial options, where the subjectively held prior beliefs play no role.

# References


Bernanke, B. S. (1983). Irreversibility, uncertainty, and cyclical investment. *The Quarterly Journal of Economics*, *98*(1), 85–106.

Bikhchandani, S., Hirshleifer, D., & Welch, I. (1992). A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of Political Economy*, *100*(5), 992–1026.

Bikhchandani, S., Segal, U., & Sharma, S. (1992). Stochastic dominance under Bayesian learning. *Journal of Economic Theory*, *56*(2), 352–377.

Bikhchandani, S. & Sharma, S. (1996). Optimal search with learning. *Journal of Economic Dynamics and Control*, *20*(1-3), 333–359.

Blackwell, D. (1953). Equivalent comparisons of experiments. *The Annals of Mathematical Statistics*, *24*(2), 265–272.

Capéraà, P. (1988). Tail ordering and asymptotic efficiency of rank tests. *The Annals of Statistics*, *16*(1), 470–478.

Chambers, C. P. & Healy, P. J. (2009). On the robustness of good news and bad news. *Ohio State University Working Paper*.

Charness, G., Karni, E., & Levin, D. (2007). Individual and group decision making under risk: An experimental study of bayesian updating and violations of first-order stochastic dominance. *Journal of Risk and Uncertainty*, *35*, 129–148.

Chetty, R. (2007). Interest rates, irreversibility, and backward-bending investment. *The Review of Economic Studies*, *74*(1), 67–91.

Cox, J. C., Ross, S. A., & Rubinstein, M. (1979). Option pricing: A simplified approach. *Journal of financial Economics*, *7*(3), 229–263.

Cukierman, A. (1980). The effects of uncertainty on investment under risk neutrality with endogenous information. *Journal of Political Economy*, *88*(3), 462–475.

Demers, M. (1991). Investment under uncertainty, irreversibility and the arrival of information over time. *The Review of Economic Studies*, *58*(2), 333–350.

Dixit, R. K. & Pindyck, R. S. (1994). *Investment under uncertainty*. Princeton University Press.

Fahmy, S., Carlos, A., Pereira, B., Proschan, F., & Shaked, M. (1982). The influence of the sample on the posterior distribution. *Communications in Statistics – Theory and Methods*, *11*(16), 1757–1768.

Ifrach, B., Maglaras, C., Scarsini, M., & Zseleva, A. (2019). Bayesian social learning from consumer reviews. *Operations Research*, *67*(5), 1209–1221.

Lehrer, E. & Wang, T. (2023). Weighted utility and optimism/pessimism: A decision-theoretic foundation of various stochastic dominance orders. *American Economic Journal: Microeconomics*, forthcoming.

Levy, H. (2006). *Stochastic dominance: Investment decision making under uncertainty*. Springer.

Milgrom, P. R. (1981). Good news and bad news: Representation theorems and applications. *The Bell Journal of Economics*, *12*(2), 380–391.

Müller, A. & Scarsini, M. (2002). Even risk-averters may love risk. *Theory and Decision*, *52*, 81–99.

Perrakis, S. (2019). *Stochastic dominance option pricing: An alternative approach to option market research*. Springer.

Pindyck, R. S. (1991). Irreversibility, uncertainty, and investment. *Journal of Economic Literature*, *29*(3), 1110–1148.

Shaked, M. & Shanthikumar, J. G. (2007). *Stochastic orders*. Springer Science & Business Media.

Shreve, S. (2004). *Stochastic calculus for finance I: The binomial asset pricing model*. Springer Science & Business Media.

Whitt, W. (1979). A note on the influence of the sample on the posterior distribution. *Journal of the American Statistical Association*, *74*(366), 424–426.

# A Appendix

## A.1 Proof of Proposition 1

In order to prove Proposition 1, we first prove the following two lemmas. They will also be useful to establish other results in the paper. Lemma 1 says that the probability of observing a length-$t$ history whose empirical frequency is at least $\frac{k}{t}$ is strictly increasing in the state $p$.

**Lemma 1.** *For any $k$, $0 \leq k \leq t$, let $\Gamma(k,t;p) := \sum_{i=k}^{t} \binom{t}{i} p^i (1-p)^{t-i}$ and let $\bar{\Gamma}(k,t;p) := \sum_{i=0}^{k} \binom{t}{i} p^i (1-p)^{t-i}$. Then $\Gamma(k,t;p)$ is increasing in $p$, and $\bar{\Gamma}(k,t;p)$ is decreasing in $p$.*

*Proof.* Note that $\Gamma(k,t;p)$ is the probability of observing a length-$t$ history with frequency no less than $\frac{k}{t}$, conditional on the true state of nature being $p$.

We prove by induction on $k$ (start with $t$, and then $t-1, ..., k+1, k$, so on so forth) that

$$\frac{\mathrm{d}\Gamma(k,t;p)}{\mathrm{d}p} = \frac{\mathrm{d}}{\mathrm{d}p}\left(\sum_{i=k}^{t} \binom{t}{i} p^i (1-p)^{t-i}\right) = \frac{t!}{(k-1)!(t-k)!}\, p^{k-1}(1-p)^{t-k} > 0. \quad (7)$$

For $k = t$, $\Gamma(t,t;p) = p^t$, so $\frac{\mathrm{d}p^t}{\mathrm{d}p} = tp^{t-1}$, which satisfies Eq. (7). Similarly, for $k = t-1$,

$$\frac{\mathrm{d}\Gamma(t-1,t;p)}{\mathrm{d}p} = \frac{\mathrm{d}(p^t + tp^{t-1}(1-p))}{\mathrm{d}p} = t(t-1)p^{t-2}(1-p),$$

so again Eq. (7) is satisfied.

Now suppose Eq. (7) is satisfied for $k+1$, namely

$$\frac{\mathrm{d}\Gamma(k+1,t;p)}{\mathrm{d}p} = \frac{t!}{k!(t-k-1)!}\, p^{k}(1-p)^{t-k-1}. \quad (8)$$

We show that Eq. (7) also holds for $k$.

Notice that

$$\frac{\mathrm{d}}{\mathrm{d}p}\binom{t}{k} p^k (1-p)^{t-k} = \frac{t!}{k!(t-k)!}\, p^{k-1}(1-p)^{t-k-1}(k-tp). \quad (9)$$

Using Eq. (8) and Eq. (9), we have

$$\begin{aligned}
\frac{\mathrm{d}\Gamma(k,t;p)}{\mathrm{d}p} &= \frac{\mathrm{d}\Gamma(k+1,t;p)}{\mathrm{d}p} + \frac{\mathrm{d}}{\mathrm{d}p}\binom{t}{k}p^k(1-p)^{t-k} \\
&= \frac{t!}{k!(t-k-1)!}\,p^k(1-p)^{t-k-1} + \frac{t!}{k!(t-k)!}\,p^{k-1}(1-p)^{t-k-1}(k-tp) \\
&= \frac{t!}{k!(t-k)!}\,p^{k-1}(1-p)^{t-k-1}\Big((t-k)p+(k-tp)\Big) \\
&= \frac{t!}{k!(t-k)!}\,p^{k-1}(1-p)^{t-k-1}k(1-p) \\
&= \frac{t!}{(k-1)!(t-k)!}\,p^{k-1}(1-p)^{t-k},
\end{aligned}$$

which is exactly Eq. (7). Since $\bar{\Gamma}(k,t;p) = 1-\Gamma(k+1,t;p)$, we conclude that $\bar{\Gamma}(k,t;p)$ is decreasing in $p$. This completes the proof for the lemma. □

The aim of the following lemma is to calculate the probability of observing those histories whose empirical frequencies belong to a given interval in the limiting case when the number of observations goes to infinity. Recall that by $\theta(h_t)$ we denote the empirical frequency of a history $h_t$.

**Lemma 2.** *Let $F$ be a prior belief (CDF) of $p$ on $[0,1]$. Consider the set of histories $H_t(p_L,p_H) := \{h_t \in \mathscr{H}_t \mid p_L \le \theta(h_t) \le p_H\}$, where $p_L$, $p_H$ are constants that satisfy $0 \le p_L < p_H \le 1$. Then as $t \to \infty$, $\mathbb{P}_F(H_t(p_L,p_H)) \to F(p_H) - F(p_L)$.*

*Proof.* We prove the assertion for the case where both $p_L$ and $p_H$ are not atoms and $p_H < 1$. The proof for the case where $p_H = 1$ is similar and is omitted. Let $B(p;h_t) := \big(p^{\theta(h_t)}(1-p)^{1-\theta(h_t)}\big)^t$ denote the probability of observing history $h_t$ when the true state is $p$, and let $B(p;H_t) := \sum_{h_t\in H_t} B(p;h_t)$ be the probability of observing a history in the set $H_t$. By the Weak Law of Large Numbers, we have $\lim_{t\to\infty} B(p;H_t) = 1$ for $p \in (p_L,p_H)$, and $\lim_{t\to\infty} B(p;H_t) = 0$ for $p < p_L$ and $p > p_H$. Since $0 \le B(p;H_t) \le 1$ for every $p$, the dominated convergence theorem implies that $\lim_{t\to\infty}\mathbb{P}_F(H_t(p_L,p_H)) = \lim_{t\to\infty}\int_0^1 B(p;H_t)\mathrm{d}F = F(p_H) - F(p_L)$. □

With Lemma 1 and Lemma 2, we are ready to prove Proposition 1.

**Proof of Proposition 1**

"Only if" direction. To show that $\{\mathbb{P}_F(k,t)\}_{k=0}^t \succeq_{\text{FOSD}} \{\mathbb{P}_G(k,t)\}_{k=0}^t$, it suffices to show that for any $k$, where $0 \le k \le t$,

$$\sum_{i=k}^{t}\mathbb{P}_F(i,t) \ge \sum_{i=k}^{t}\mathbb{P}_G(i,t).$$

Note that, by Eq. (1),

$$\begin{aligned}\sum_{i=k}^{t}\mathbb{P}_F(i,t) &= \int_0^1 \sum_{i=k}^{t}\binom{t}{i}p^i(1-p)^{t-i}\mathrm{d}F \\ &= \int_0^1 \Gamma(k,t;p)\mathrm{d}F \\ &= \mathbb{E}_F(\Gamma(k,t;p)).\end{aligned}$$

Similarly, $\sum_{i=k}^{t}\mathbb{P}_G(i,t)=\mathbb{E}_G(\Gamma(k,t;p))$.

Since for any $(k,t)$, Lemma 1 tells us that $\Gamma(k,t;p)$ is increasing in $p$, the assumption that $F \succeq_{\text{FOSD}} G$ implies that $\mathbb{E}_F(\Gamma(k,t;p)) \geq \mathbb{E}_G(\Gamma(k,t;p))$, hence $\sum_{i=k}^{t}\mathbb{P}_F(i,t) \geq \sum_{i=k}^{t}\mathbb{P}_G(i,t)$ for $0 \leq k \leq t$. This completes the proof of the "only if" direction.

"If" direction. Suppose that $F \succeq_{\text{FOSD}} G$ does not hold. Then there exists a $\hat{p} \in (0,1)$ such that $F(\hat{p}) > G(\hat{p})$. Consider the set $H_t(\hat{p}) := \{h_t \in \mathscr{H}_t \mid 0 \leq \theta(h_t) \leq \hat{p}\}$ of length-$t$ histories. It follows from $\{\mathbb{P}_F(k,t)\}_{k=0}^{t} \succeq_{\text{FOSD}} \{\mathbb{P}_G(k,t)\}_{k=0}^{t}$ that $\mathbb{P}_F(H_t(\hat{p})) \leq \mathbb{P}_G(H_t(\hat{p}))$ for any $t$. By Lemma 2, as $t \to \infty$, $\mathbb{P}_F(H_t(\hat{p})) \to F(\hat{p})$ and $\mathbb{P}_G(H_t(\hat{p})) \to G(\hat{p})$, hence in the limit, it must be true that $F(\hat{p}) \leq G(\hat{p})$, which contradicts $F(\hat{p}) > G(\hat{p})$. This completes the proof for the "if" direction.

## A.2 Proof of Theorem 1 and Supplementary Results

*Proof.* Let us focus on item (i) of Theorem 1. The proof of (ii) is similar. To show the "only if" direction, recall from Proposition 1 that $F \succeq_{\text{FOSD}} G$ implies that $\{\mathbb{P}_F(k,T)\}_{k=0}^{T} \succeq_{\text{FOSD}} \{\mathbb{P}_G(k,T)\}_{k=0}^{T}$. Since $v(k,T)$ is monotone in $k$, we conclude from Eq. (2) that $V_F(T,\bar{S},S(k,T)) \geq V_G(T,\bar{S},S(k,T))$ for every European call option.

To show the "if" direction, according to Proposition 1, it suffices to show that the greater value under $F$ for very European call option implies that for every period $t$, $\{\mathbb{P}_F(k,t)\}_{k=0}^{t} \succeq_{\text{FOSD}} \{\mathbb{P}_G(k,t)\}_{k=0}^{t}$. To this end, consider $t$ different European call options with $T=t$ and $S^i(k,t)=\begin{cases}1+\bar{S}, & \text{if } i \leq k \leq t,\\ \bar{S}, & \text{if } 0 \leq k < i,\end{cases}$ for $i=1,\ldots,t$. The ex post value of the $i$-th option is $v^i(k,t,S^i(k,t))=\max\{0,S^i(k,t)-\bar{S}\}=\begin{cases}1, & \text{if } i \leq k \leq t,\\ 0, & \text{if } 0 \leq k < i,\end{cases}$ which is increasing in $k$. By Eq. (2), the assumption that $V_F^i(t,\bar{S},S^i(k,t)) \geq V_G^i(t,\bar{S},S^i(k,t))$ for $i=1,\ldots,t$ implies that $\sum_{k=i}^{t}\mathbb{P}_F(k,t) \geq \sum_{k=i}^{t}\mathbb{P}_G(k,t)$, for $i=1,\ldots,t$, which shows that $\{\mathbb{P}_F(k,t)\}_{k=0}^{t} \succeq_{\text{FOSD}} \{\mathbb{P}_G(k,t)\}_{k=0}^{t}$. □

In the proof of the "if" direction, in order to ensure that $\{\mathbb{P}_F(k,t)\}_{k=0}^{t} \succeq_{\text{FOSD}} \{\mathbb{P}_G(k,t)\}_{k=0}^{t}$, we considered $t$ special European options (increasing functions $S^i(k,t)$). One may wonder whether we can consider a general set of increasing functions instead.

In what follows, we show that to check whether $\{\mathbb{P}_F(k,t)\}_{k=0}^t \succeq_{\text{FOSD}} \{\mathbb{P}_G(k,t)\}_{k=0}^t$ holds, it suffices to consider $t$ increasing payoff functions that satisfy certain mild conditions.

Let $f = (f_0, \ldots, f_t)$, $g = (g_0, \ldots, g_t)$ be two probability distributions over $t+1$ outcomes, and let $F$, $G$ be their CDFs. Consider any $t$ different increasing payoff functions $R^1, \ldots, R^t$ defined on $t+1$ outcomes.

**Definition 1.** *Consider an increasing payoff function $R$ over $t+1$ outcomes: $R(i) = r_i$, $i = 0, \ldots, t$, where $r_i \geq r_j$ whenever $i > j$. We call $\Delta R := (r_i - r_{i-1})_{i=1}^t$ the* increment vector *of $R$.*

Since $R$ is increasing, the $t$ elements of $\Delta R$ are all non-negative.

**Lemma 3.** *Consider two probability mass functions $f = (f_0, ..., f_t)$, $g = (g_0, ..., g_t)$ over $t+1$ outcomes and let $F$, $G$ be their* CDFs. *To determine whether $F \succeq_{\text{FOSD}} G$, it suffices to consider any $t$ monotone payoff functions over the $t+1$ outcomes with linearly independent increment vectors.*

*Proof.* Note that for a given increasing payoff function $R^i$, its expectation under $F$ (similarly under $G$) can be written as follows:

$$\begin{aligned}
\mathbb{E}_F(R^i) &= \sum_{j=0}^t f_j r_j^i \\
&= \left(\sum_{j=0}^t f_j\right) r_0^i + \left(\sum_{j=1}^t f_j\right)(r_1^i - r_0^i) + \cdots + \left(\sum_{j=\ell}^t f_j\right)(r_\ell^i - r_{\ell-1}^i) \\
&\quad + \cdots + f_t(r_t^i - r_{t-1}^i) \\
&= r_0^i + \sum_{\ell=1}^t (r_\ell^i - r_{\ell-1}^i)\left(\sum_{j=\ell}^t f_j\right).
\end{aligned}$$

Recall that, by definition, $\Delta R^i = \left(r_\ell^i - r_{\ell-1}^i\right)_{\ell=1}^t$ is the increment vector of $R^i$. Let $\bar{F}_\ell = \sum_{j=\ell}^t f_j$ be the survival function of $F$ at $\ell$, $\ell = 0, ..., t$ (similarly for $\bar{G}_\ell$), and let $\Delta(\bar{F} - \bar{G}) := \left(\bar{F}_j - \bar{G}_j\right)_{j=1}^t$ be the vector or differences between the two survival functions. It follows that

$$\mathbb{E}_F(R^i) - \mathbb{E}_G(R^i) = \sum_{\ell=1}^t (r_\ell^i - r_{\ell-1}^i)\left(\sum_{j=\ell}^t f_j - \sum_{j=\ell}^t g_j\right) = \Delta R^i \cdot \Delta(\bar{F} - \bar{G}). \tag{10}$$

Hence, so long as the $t$ increment vectors $\Delta R^1, ..., \Delta R^t$ are linearly independent, one can uniquely determine the vector $\Delta(\bar{F} - \bar{G})$. If all its elements are non-negative, we conclude that $F \succeq_{\text{FOSD}} G$; otherwise, $F \succeq_{\text{FOSD}} G$ does not hold. □

The following lemma is concerned with the payoff functions (asset price functions) that are widely used in the classical tree model for option pricing. It establishes the linear independence among increment vectors associated with these payoff functions. The proof is straightforward and hence omitted.

**Lemma 4.** *Consider $t$ increasing payoff functions $R^i(k,t) = S_0\lambda_i^{2k-t}$, $k = 0, ..., t$, $i = 1, ..., t$, over $t+1$ outcomes $(0,t), ..., (t,t)$, where $1 < \lambda_1 < \lambda_2 < \cdots < \lambda_t$. Then the increment vectors $\Delta R^1,\ \Delta R^2, ..., \Delta R^t$ are linearly independent.*

**Remark.** If we have $t$ increasing payoff functions $R^1, ..., R^t$ such that $\mathbb{E}_F(R^i) - \mathbb{E}_G(R^i) \geq 0$ for every $i$, then we cannot draw a conclusion whether $F \succeq_{\text{FOSD}} G$ holds or not. To see the reason, suppose there are 3 outcomes and consider two payoff functions $R^1,\ R^2$ with independent increment vectors $\Delta R^1,\ \Delta R^2,\ \Delta R^1 = (r_1^1 - r_0^1, r_2^1 - r_1^1)$, $\Delta R^2 = (r_1^2 - r_0^2, r_2^2 - r_1^2)$. Let $c^i := \mathbb{E}_F(R^i) - \mathbb{E}_G(R^i)$, $i = 1, 2$. Each $c^i$ is nonnegative, but we do not know its exact value. Then, according to Eq. (10),

$$\begin{pmatrix} r_1^1 - r_0^1 & r_2^1 - r_1^1 \\ r_1^2 - r_0^2 & r_2^2 - r_1^2 \end{pmatrix} \begin{pmatrix} \bar{F}_1 - \bar{G}_1 \\ \bar{F}_2 - \bar{G}_2 \end{pmatrix} = \begin{pmatrix} c^1 \\ c^2 \end{pmatrix} \geq \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

or equivalently,

$$(\bar{F}_1 - \bar{G}_1)\begin{pmatrix} r_1^1 - r_0^1 \\ r_1^2 - r_0^2 \end{pmatrix} + (\bar{F}_2 - \bar{G}_2)\begin{pmatrix} r_2^1 - r_1^1 \\ r_2^2 - r_1^2 \end{pmatrix} = \begin{pmatrix} c^1 \\ c^2 \end{pmatrix} \geq \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

To ensure $\bar{F}_1 - \bar{G}_1 \geq 0$ and $\bar{F}_2 - \bar{G}_2 \geq 0$, the vector $\binom{c^1}{c^2}$ must lie in the cone of the two column vectors $\binom{r_1^1 - r_0^1}{r_1^2 - r_0^2}$ and $\binom{r_2^1 - r_1^1}{r_2^2 - r_1^2}$. But since $c^1,\ c^2$ could be any non-negative number, unless $\binom{r_1^1 - r_0^1}{r_1^2 - r_0^2}$ and $\binom{r_2^1 - r_1^1}{r_2^2 - r_1^2}$ coincide with the two axes (which is the case for the special choice of payoff functions in the proof of "if" part of Theorem 1), we cannot guarantee that $\bar{F}_1 - \bar{G}_1 \geq 0$ and $\bar{F}_2 - \bar{G}_2 \geq 0$. Figure 2 illustrates such a case.

## A.3 Proof of Proposition 3

**1 ⇒ 2.** By condition 1 of Proposition 2, it suffices to show that for any increasing function $u$ and non-negative, increasing function $w$ integrable w.r.t. $F$ and $G$ (the case when at least one of the denominators is 0 is straightforward)

$$\frac{\mathbb{E}_{F|H}(u(p)w(p))}{\mathbb{E}_{F|H}(w(p))} \geq \frac{\mathbb{E}_{G|H}(u(p)w(p))}{\mathbb{E}_{G|H}(w(p))}.$$

By assumption, the probability of observing $H$ is increasing in $p$ (i.e., $B(p; H)$ is increasing in $p$, where $B(p; H) = \sum_{h \in H} B(p; h)$). Hence, for any non-negative increasing

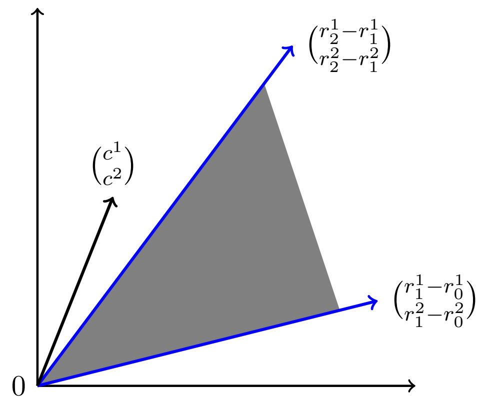


Figure 2: Greater value of European options may not imply FOSD.

function $w(p)$, the function $w(p)B(p;H)$ is non-negative, increasing and integrable. By the Bayes rule,

$$[F|H](p) = \frac{\mathbb{E}_F(B(p;H)\mathbf{1}_{[0,p]})}{\mathbb{E}_F(B(p;H))}.$$

It follows from $F \succeq_{\text{HR}} G$ that

$$\begin{aligned} \frac{\mathbb{E}_{F|H}(u(p)w(p)))}{\mathbb{E}_{F|H}(w(p))} &= \frac{\mathbb{E}_F\left(u(p)w(p)B(p;H)\right)}{\mathbb{E}_F\left(w(p)B(p;H)\right)} \\ &\geq \frac{\mathbb{E}_G\left(u(p)w(p)B(p;H)\right)}{\mathbb{E}_G\left(w(p)B(p;H)\right)} \\ &= \frac{\mathbb{E}_{G|H}(u(p)w(p)))}{\mathbb{E}_{G|H}(w(p))}. \end{aligned}$$

This is what we set out to prove. More generally, using the same argument, we can show that $F \succeq_{\text{HR}} G$ implies that for any event whose probability is increasing in the state of nature $p$, the posterior beliefs conditional on this event preserve the HRD relation.

**2 $\Rightarrow$ 3.** This is trivial, since, by Lemma 1, $B(p; H_t^+(\theta))$ is increasing in $p$ for any $t$ and any $\theta \in (0,1)$. Moreover, both HRD and RHRD imply FOSD.

**3 $\Rightarrow$ 1.** Suppose $F \succeq_{\text{HR}} G$ does not hold. By condition 2 of Proposition 2, there exists $p_H > p_L$ in $[0,1]$ such that $[1-F(p_H)][1-G(p_L)] < [1-G(p_H)][1-F(p_L)]$. Consequently, $1-F(p_L) > 0$, $1-G(p_H) > 0$, and $p_H < 1$, hence

$$\frac{1-F(p_H)}{1-F(p_L)} < \frac{1-G(p_H)}{1-G(p_L)}. \tag{11}$$

Consider the sets of histories $H_t^+(p_H) := \{h_t \in \mathscr{H}_t | \theta(h_t) > p_H\}$ and $H_t^+(p_L) := \{h_t \in \mathscr{H}_t | \theta(h_t) > p_L\}$. By Lemma 2, $\lim_{t\to\infty} \mathbb{P}_F(H_t^+(p_L)) = 1 - F(p_L)$, $\lim_{t\to\infty} \mathbb{P}_G(H_t^+(p_L)) =$

$1-G(p_L)$, $\lim_{t\to\infty}\int_0^1 B(p;H_t^+(p_L))\mathbf{1}_{(p_H,1]}\mathrm{d}F = \int_0^1 \mathbf{1}_{(p_L,1]}\mathbf{1}_{(p_H,1]}\mathrm{d}F = 1-F(p_H)$, and $\lim_{t\to\infty}\int_0^1 B(p;H_t^+(p_L))\mathbf{1}_{(p_H,1]}\mathrm{d}G = 1-G(p_H)$. It follows that

$$\begin{aligned}\lim_{t\to\infty} 1-\left[F|H_t^+(p_L)\right](p_H) &= \lim_{t\to\infty}\mathbb{E}_F(\mathbf{1}_{(p_H,1]}|H_t^+(p_L)) \\ &= \lim_{t\to\infty}\frac{\int_0^1 B(p;H_t^+(p_L))\mathbf{1}_{(p_H,1]}\mathrm{d}F}{\mathbb{P}_F(H_t^+(p_L))} \\ &= \frac{1-F(p_H)}{1-F(p_L)},\end{aligned}$$

and similarly,

$$\lim_{t\to\infty} 1-\left[G|H_t^+(p_L)\right](p_H) = \frac{1-G(p_H)}{1-G(p_L)}.$$

By assumption, $F|H_t^+(p_L) \succeq_{\text{FOSD}} G|H_t^+(p_L)$, hence for every $t$, $1-\left[F|H_t^+(p_L)\right](p_H) \geq 1-\left[G|H_t^+(p_L)\right](p_H)$, which implies that $\frac{1-F(p_H)}{1-F(p_L)} \geq \frac{1-G(p_H)}{1-G(p_L)}$, contradicting Eq. (11).

**1 ⇒ 4.** We first establish the following lemma.

**Lemma 5.** *Suppose a set of length-$t$ histories has $\alpha_i$ histories with empirical frequency $\frac{k_i}{t}$, $i=1,...,N$, where $0\leq k_i<k_j\leq t$ whenever $i<j$, and each $\alpha_i$ is a non-negative integer. Let $n$ be any number between 1 and $N$. Then the expression*

$$\frac{\sum_{i>n}\alpha_i p^{k_i}(1-p)^{t-k_i}}{\sum_{j\leq n}\alpha_j p^{k_j}(1-p)^{t-k_j}} \tag{12}$$

*is strictly increasing in $p$, for $p\in(0,1)$.*

*Proof.* Divide the numerator and the denominator in Eq. (12) by $p^{k_n}(1-p)^{t-k_n}$. Then (12) becomes

$$\frac{\sum_{i>n}\alpha_i\left(\frac{p}{1-p}\right)^{k_i-k_n}}{\sum_{j\leq n}\alpha_j\left(\frac{p}{1-p}\right)^{k_j-k_n}}. \tag{13}$$

Since $k_i>k_n$ for all $i>n$ and $k_j<k_n$ for all $j<n$, and since $\frac{p}{1-p}$ is strictly increasing in $p$, we conclude that all terms in the numerator of (13) are strictly increasing in $p$, and all terms in the denominator of (13) are strictly decreasing in $p$, hence (12) is strictly increasing in $p$. □

Take any $\theta\in[0,1]$ and consider any history $h_t\in H_t^+(\theta)$. Then $h_t$ partitions $H_t^+(\theta)$ into two disjoint subsets, $H_t^{\geq}$ and $H_t^{<}$, where $H_t^{\geq}$ contains those histories in $H_t^+(\theta)$ whose empirical frequencies are not smaller than that of $h_t$, and $H_t^{<}$ is the collection of those histories in $H_t^+(\theta)$ with strictly smaller frequencies. Let $B(p;H_t^{\geq})$, $B(p;H_t^{<})$

be the probabilities of observing a history in the sets $H_t^{\geq}$, $H_t^{<}$, respectively, when the true state is $p$. It suffices to show that $\frac{\mathbb{E}_F(B(p;H_t^{\geq}))}{\mathbb{E}_F(B(p;H_t^+(\theta)))} \geq \frac{\mathbb{E}_G(B(p;H_t^{\geq}))}{\mathbb{E}_G(B(p;H_t^+(\theta)))}$. By Lemma 1, $B(p;H_t^+(\theta))$ and $B(p;H_t^{\geq})$ are increasing in $p$, and by Lemma 5, $\frac{B(p;H_t^{\geq})}{B(p;H_t^{<})}$ is strictly increasing in $p$, hence $\frac{B(p;H_t^{\geq})}{B(p;H_t^+(\theta))} = \frac{B(p;H_t^{\geq})}{B(p;H_t^{\geq})+B(p;H_t^{<})}$ is strictly increasing in $p$. It follows from the first equivalent condition of Proposition 2 that

$$\begin{aligned}
\frac{\mathbb{E}_F(B(p;H_t^{\geq}))}{\mathbb{E}_F(B(p;H_t^+(\theta)))} &= \frac{\mathbb{E}_F\left(B(p;H_t^+(\theta))\frac{B(p;H_t^{\geq})}{B(p;H_t^+(\theta))}\right)}{\mathbb{E}_F(B(p;H_t^+(\theta)))} \\
&\geq \frac{\mathbb{E}_G\left(B(p;H_t^+(\theta))\frac{B(p;H_t^{\geq})}{B(p;H_t^+(\theta))}\right)}{\mathbb{E}_G(B(p;H_t^+(\theta)))} \\
&= \frac{\mathbb{E}_G(B(p;H_t^{\geq}))}{\mathbb{E}_G(B(p;H_t^+(\theta)))}.
\end{aligned}$$

This shows that the probability distribution over histories in $H_t^+(\theta)$ under $F$ FOSD that under $G$.

**4 ⇒ 1.** Suppose $F \not\succeq_{\mathrm{HR}} G$. Then there exits $p_H > p_L$ in $[0,1]$ such that Eq. (11) holds. Let $H_t^L := \{h_t \in \mathscr{H}_t | \theta(h_t) > p_L\}$, and let $H_t^H := \{h_t \in \mathscr{H}_t | \theta(h_t) > p_H\}$. By Lemma 2, we know that as $t \to \infty$, $\mathbb{P}_F(H_t^L) \to 1 - F(p_L)$, $\mathbb{P}_F(H_t^H) \to 1 - F(p_H)$, and similarly, $\mathbb{P}_G(H_t^L) \to 1 - G(p_L)$, $\mathbb{P}_G(H_t^H) \to 1 - G(p_H)$.

By assumption, the probability distribution over $H_t^L$ under $F$ FOSD the probability distribution over $H_t^L$ under $G$, for every $t$, and so $\frac{\mathbb{P}_F(H_t^H)}{\mathbb{P}_F(H_t^L)} \geq \frac{\mathbb{P}_G(H_t^H)}{\mathbb{P}_G(H_t^L)}$. In the limit, we have $\frac{1-F(p_H)}{1-F(p_L)} \geq \frac{1-G(p_H)}{1-G(p_L)}$, which contradicts Eq. (11).

## A.4 Proof of Proposition 4

The proof of the only if direction is essentially the same as the proof of "1 ⇒ 2" of Proposition 3. We only need to replace $B(p;H_t^+(\theta))$ with $\pi(E|p)$ and note that $\pi(E|p)$ is increasing (resp., decreasing) in $p$ by assumption.

For the if direction, regard the conditional probability distribution over length-$t$ histories given state $p$ as a Blackwell experiment. The desired result follows from the proof of "3 ⇒ 1" of Proposition 3.

## A.5 Proof of Theorem 2

For the "only if" direction, note that for any European call option there exists a cutoff $\theta$ such that the option is exercised in the termination date when the empirical frequency is above $\theta$. Let $H_T^+(\theta) = \{h_T \in \mathscr{H}_T \mid \theta(h_T) \geq \theta\}$ be the set of length-$T$ histories whose frequencies are greater than $\theta$. It follows from Proposition 3 (the equivalence between conditions 1 and 4) that the probability distribution over the frequencies in $H_T^+(\theta)$ under $F$ first-order stochastically dominates that under $G$. This together with the fact that $S(k,T)$ is increasing in $k$ yields the desired result.

To show the "if" direction, we note that by the equivalence of conditions 1 and 4 of Proposition 3, it suffices to show that for any $\theta \in (0,1)$ and any $T$, $\{\mathbb{P}_F(H_T^+(\theta))\} \succeq_{\text{FOSD}} \{\mathbb{P}_G(H_T^+(\theta))\}$. Suppose this set of histories contains $n+1$ different outcomes (i.e., $n+1$ frequencies) $(T,T), ..., (T-n,T)$. Consider the following $n$ European call options with termination date $T$ and

$$S^i(k,T) = \begin{cases} \bar{S}+1, & \text{if } T-i \leq k \leq T, \\ \bar{S}, & \text{if } T-n \leq k < T-i, \\ S(k,T), & \text{if } 0 \leq k < T-n \end{cases}$$

for $i = 0, ..., n-1$. Here $S^i(k,T)$ is increasing in $k$ and is strictly less than $\bar{S}$ (hence the option will not be exercised) for $0 \leq k < T-n$. It follows from the assumption that the value of every option $i$ conditional on it is being exercised (i.e., $T-n \leq k \leq T$) is greater under $F$ than under $G$ that $\{\mathbb{P}_F(H_T^+(\theta))\} \succeq_{\text{FOSD}} \{\mathbb{P}_G(H_T^+(\theta))\}$. For European put options, the proof is similar.

## A.6 Proof of Proposition 6

**1 $\Rightarrow$ 2.** According to the first equivalent condition of Proposition 5 (Eq. (4)), to show that $F|H \succeq_{\text{LR}} G|H$ for any set of histories $H$, it suffices to show that for any increasing function $u$ and any non-negative function $w$ integrable w.r.t. $F$ and $G$,

$$\frac{\mathbb{E}_{F|H}(u(p)w(p))}{\mathbb{E}_{F|H}(w(p))} \geq \frac{\mathbb{E}_{G|H}(u(p)w(p))}{\mathbb{E}_{G|H}(w(p))}.$$

(In case one of the denominators is 0, Eq. (4) holds trivially.) It follows from $F \succeq_{\text{LR}} G$ that

$$\begin{aligned} \frac{\mathbb{E}_{F|H}(u(p)w(p))}{\mathbb{E}_{F|H}(w(p))} &= \frac{\mathbb{E}_F\left(u(p)w(p)B(p;H)\right)}{\mathbb{E}_F\left(w(p)B(p;H)\right)} \\ &\geq \frac{\mathbb{E}_G\left(u(p)w(p)B(p;H)\right)}{\mathbb{E}_G\left(w(p)B(p;H)\right)} \\ &= \frac{\mathbb{E}_{G|H}(u(p)w(p))}{\mathbb{E}_{G|H}(w(p))}. \end{aligned}$$

This is what we set out to prove.

**2 ⇒ 3.** This is trivial, since LRD implies FOSD.

**3 ⇒ 1.** According to the second equivalent condition of Proposition 5, it suffices to show that for each $p_3 < p_2 < p_1$ in $[0, 1]$,

$$[F(p_1) - F(p_2)][G(p_1) - G(p_3)] \geq [G(p_1) - G(p_2)][F(p_1) - F(p_3)]. \tag{14}$$

We focus on the case in which $G(p_1) - G(p_3) > 0$ and $F(p_1) - F(p_3) > 0$. If one of these differences is zero, say $G(p_1) - G(p_3) = 0$, then $G(p_1) - G(p_2) = 0$ and Eq. (14) holds trivially. Therefore, we show that

$$\frac{F(p_1) - F(p_2)}{F(p_1) - F(p_3)} \geq \frac{G(p_1) - G(p_2)}{G(p_1) - G(p_3)}. \tag{15}$$

First, suppose that $p_1 < 1$ (the proof for the case $p_1 = 1$ is the same and omitted). Since $F$ and $G$ are right-continuous and each has at most countably many atoms, we can assume w.l.o.g. that there are no atoms (neither of $F$, nor of $G$) at $p_3$, $p_2$ and $p_1$. Indeed, suppose that $p_1 < 1$ and $p_1$ is an atom. One can find a point $q_1 > p_1$ which is not an atom of $F$ or $G$ (because there are at most countably many atoms). Moreover, since every CDF is right-continuous, $q_1$ can be chosen to be very close to $p_1$ (from the right), so that $F(q_1)$ and $G(q_1)$ are close to $F(p_1)$ and $G(p_1)$ to the extent that the inequality in Eq. (15) remains valid when we replace $p_1$ by $q_1$. For the same reason, w.l.o.g., we can assume that $p_3$ and $p_2$ are not atoms.

Consider the set of equal-length histories $H_t := \{h_t \in \mathscr{H}_t \mid p_3 < \theta(h_t) \leq p_1\}$. By Lemma 2, $\lim_{t\to\infty} \mathbb{P}_F(H_t) = F(p_1) - F(p_3)$ and $\lim_{t\to\infty} \int_0^1 B(p; H_t)\mathbf{1}_{(p_2,1]}\mathrm{d}F = \int_0^1 \mathbf{1}_{(p_3,p_1]}\mathbf{1}_{(p_2,1]}\mathrm{d}F = F(p_1) - F(p_2)$. It follows that

$$\begin{aligned}
\lim_{t\to\infty}(1 - [F|H_t](p_2)) &= \lim_{t\to\infty} \mathbb{E}_F(\mathbf{1}_{(p_2,1]}|H_t) \\
&= \lim_{t\to\infty} \frac{\int_0^1 B(p; H_t)\mathbf{1}_{(p_2,1]}\mathrm{d}F}{\mathbb{P}_F(H_t)} \\
&= \frac{F(p_1) - F(p_2)}{F(p_1) - F(p_3)}.
\end{aligned}$$

Similarly,

$$\lim_{t\to\infty}(1 - [G|H_t](p_2)) = \frac{G(p_1) - G(p_2)}{G(p_1) - G(p_3)}.$$

Since $F|H_t \succeq_{\mathrm{FOSD}} G|H_t$, we can deduce that Eq. (15) holds.

Now consider the case in which $p_3 = 0$. Using the same argument, we can show that for each $p_2 < p_1$ in $(0, 1]$, $[F(p_1) - F(p_2)]G(p_1) \geq [G(p_1) - G(p_2)]F(p_1)$. This completes the proof that condition 3 implies that $F \succeq_{\mathrm{LR}} G$.

**1 ⇒ 4.** Fix a given set of histories $H_t$. Consider a history $h_t \in H_t$. Let $\theta(h_t)$ be the empirical frequency of $h_t$. The history $h_t$ partitions $H_t$ into two disjoint subsets, $H_t^+$ and $H_t^-$, where $H_t^+$ contains those histories in $H_t$ whose frequencies are no smaller than that of $h_t$, and $H_t^-$ is the collection of those histories in $H_t$ with strictly smaller frequencies. Let $B(p; H_t^+)$, $B(p; H_t^-)$ be the probabilities of observing a history in the set $H_t^+$, $H_t^-$ respectively, when the true state is $p$. It suffices to show that $\frac{\mathbb{E}_F(B(p;H_t^+))}{\mathbb{E}_F(B(p;H_t^-))} \geq \frac{\mathbb{E}_G(B(p;H_t^+))}{\mathbb{E}_G(B(p;H_t^-))}$. By Lemma 5, $\frac{B(p;H_t^+)}{B(p;H_t^-)}$ is strictly increasing in $p$. Hence, it follows from the first equivalent condition of Proposition 5 (Eq. (4)) that

$$\begin{aligned}
\frac{\mathbb{E}_F(B(p; H_t^+))}{\mathbb{E}_F(B(p; H_t^-))} &= \frac{\mathbb{E}_F\left(B(p; H_t^-)\frac{B(p;H_t^+)}{B(p;H_t^-)}\right)}{\mathbb{E}_F(B(p; H_t^-))} \\
&\geq \frac{\mathbb{E}_G\left(B(p; H_t^-)\frac{B(p;H_t^+)}{B(p;H_t^-)}\right)}{\mathbb{E}_G(B(p; H_t^-))} \\
&= \frac{\mathbb{E}_G(B(p; H_t^+))}{\mathbb{E}_G(B(p; H_t^-))}.
\end{aligned}$$

This shows that the probability distribution over histories in $H_t$ under $F$ FOSD that under $G$.

**4 ⇒ 1.** Suppose $F \not\geq_{\text{LR}} G$. Then by Eq. (5) of Proposition 5, there exist $p_1, p_2, p_3 \in [0, 1]$ with $p_1 > p_2 > p_3$ such that

$$\frac{F(p_2) - F(p_3)}{F(p_1) - F(p_3)} > \frac{G(p_2) - G(p_3)}{G(p_1) - G(p_3)}. \tag{16}$$

As in the proof of "3 ⇒ 1", w.l.o.g., one can assume that $p_1,\ p_2,\ p_3$ are not atoms (or $p_1 = 1$).

For any $p,\ q$ with $0 \leq p < q \leq 1$, consider the set of histories $H_t(p, q) := \{h_t \in \mathscr{H}_t \mid p < \theta(h_t) \leq q\}$, where $\theta(h_t)$ is the empirical frequency of $h_t$. By Lemma 2, $\mathbb{P}_F(H_t(p_3, p_2)) \to F(p_2) - F(p_3)$ and $\mathbb{P}_F(H_t(p_3, p_1)) \to F(p_1) - F(p_3)$ as $t \to \infty$. Therefore, as $t \to \infty$,

$$\frac{\mathbb{P}_F(H_t(p_3, p_2))}{\mathbb{P}_F(H_t(p_3, p_1))} \to \frac{F(p_2) - F(p_3)}{F(p_1) - F(p_3)}.$$

Similarly, under prior belief $G$, as $t \to \infty$,

$$\frac{\mathbb{P}_G(H_t(p_3, p_2))}{\mathbb{P}_G(H_t(p_3, p_1))} \to \frac{G(p_2) - G(p_3)}{G(p_1) - G(p_3)}.$$

By condition 4, the probability distribution on the set of histories $H_t(p_3, p_1)$ under $F$ FOSD that under $G$. Consequently, $\frac{\mathbb{P}_F(H_t(p_3,p_2))}{\mathbb{P}_F(H_t(p_3,p_1))} \leq \frac{\mathbb{P}_G(H_t(p_3,p_2))}{\mathbb{P}_G(H_t(p_3,p_1))}$ for any $t$. Hence, in the limit,

$$\frac{F(p_2) - F(p_3)}{F(p_1) - F(p_3)} \leq \frac{G(p_2) - G(p_3)}{G(p_1) - G(p_3)},$$

which contradicts Eq. (16). This completes the proof of Propositoin 6.

## A.7 Proof of Proposition 8

We first establish the following result that compares the option value conditional on a single history $h_t$.

**Lemma 6.** *Suppose $F$ and $G$ are two non-trivial priors that satisfy $F \succeq_{\text{LR}} G$. Then the value of every American call option conditional on every history $h_t$ is greater under $F$ than under $G$, i.e., $V_F^T(k,t) \geq V_G^T(k,t)$ for every $(k,t)$, $t \leq T$.*

*Proof.* Consider an American option characterized by $(T, \bar{S}, S(k,t))$. We proceed by induction. For the last period $T$, $V_F^T(k,T) = V_G^T(k,T) = \max\{0, S(k,T) - \bar{S}\}$ for any $0 \leq k \leq T$, so the desired result holds trivially. Now suppose the result holds for $t+1$, i.e., $V_F^T(k,t+1) \geq V_G^T(k,t+1)$ for any $0 \leq k \leq t+1$. In period $t$, suppose we observe a history $h_t$ with outcome $(k,t)$. By Proposition 6, the posteriors satisfy $F|h_t \succeq_{\text{FOSD}} G|h_t$. By Proposition 1, conditional on $(k,t)$, the probability distribution over the two nodes $(k+1,t+1)$ and $(k,t+1)$ under $F$ first-order stochastically dominates that under $G$. Hence, $\mathbb{P}_F(U|k,t) \geq \mathbb{P}_G(U|k,t)$. Since $V_F^T(k+1,t+1) \geq V_G^T(k+1,t+1)$ and $V_F^T(k,t+1) \geq V_G^T(k,t+1)$, we conclude that $V_F^T(k,t) \geq V_G^T(k,t)$ for any $(k,t)$. □

When $F \succeq_{\text{LR}} G$, by Lemma 6, the expected continuation value conditional on every history under $F$ is greater than that under $G$. We conclude that the set of histories that lead to the exercise of an American call option is smaller under a likelihood-ratio dominating prior.

## A.8 Proof of Theorem 3

To show the "only if" direction, recall from Proposition 6 (the equivalence of condition 1 and condition 4) that $F \succeq_{\text{LR}} G$ implies that the probability distribution over $H_t$ under $F$ FOSD that under $G$. Specifically, the probability distribution over histories in $H_t$ is

$$\mathbb{P}_F(h_t|H_t) = \frac{\mathbb{E}_F(B(p;h_t))}{\mathbb{E}_F(B(p;H_t))},$$

where $B(p;h_t) = \left(p^{\theta(h_t)}(1-p)^{1-\theta(h_t)}\right)^t$ is the probability of observing history $h_t$ when the true state is $p$, and $B(p;H_t) = \sum_{h_t \in H_t} B(p;h_t)$. The expected option value conditional on $H_t$ under prior belief $F$ (similarly, under $G$) is

$$\sum_{h_t \in H_t} \mathbb{P}_F(h_t|H_t) V_F(\theta(h_t)t, t).$$

The desired result follows from Lemma 6 and that $\{\mathbb{P}_F(h_t|H_t)\} \succeq_{\text{FOSD}} \{\mathbb{P}_G(h_t|H_t)\}$.

For the "if" direction, according to the proof of Proposition 6 ($4 \Rightarrow 1$), it suffices to consider the sets of histories of the form $H_t(\underline{\theta}, \bar{\theta}) := \{h_t \in \mathscr{H}_t | \underline{\theta} \leq \theta(h_t) \leq \bar{\theta}\}$ for any $0 \leq \underline{\theta} < \bar{\theta} \leq 1$ and any $t$, such that the probability distribution over each set $H_t(\underline{\theta}, \bar{\theta})$ under $F$ FOSD that under $G$.

Fix such a set $H_t(\underline{\theta}, \bar{\theta})$ and suppose that it has $n+1$ different outcomes $(k_0, t), (k_0 + 1, t), ..., (k_0 + n, t)$. Consider the following $n$ American call options with $T = t$ and

$$S^i(k,t) = \begin{cases} S(k,t), & \text{if } k_0 + n < k \leq t, \\ \bar{S} + 1, & \text{if } k_0 + i \leq k \leq k_0 + n, \\ \bar{S}, & \text{if } k_0 \leq k < k_0 + i, \\ S(k,t), & \text{if } 0 \leq k < k_0, \end{cases}$$

for $i = 1, 2, ..., n$, where $S(k,t)$ is increasing in $k$ and is larger than $\bar{S} + 1$ for $k_0 + n < k \leq t$, and smaller than $\bar{S}$ for $0 \leq k < k_0$. For $t' < t$, choose $S^i(k, t')$ such that

$$\begin{aligned} S^i(k,t') - \bar{S} \leq \delta \min \big\{ & \mathbb{P}_F(U|k,t') V_F^t(k+1, t'+1) + \mathbb{P}_F(D|k,t') V_F^t(k, t'+1), \\ & \mathbb{P}_G(U|k,t') V_G^t(k+1, t'+1) + \mathbb{P}_G(D|k,t') V_G^t(k, t'+1) \big\}, \end{aligned}$$

where $V_j^t(k+1, t'+1)$, $V_j^t(k, t'+1)$, $j = F, G$, are obtained recursively backward via Eq. (6). Hence, for each such American call option, it is optimal to hold the option until the termination date $T = t$ under both $F$ and $G$, and exercise it when $k \geq k_0$. It follows from the assumption that the expected value of each option conditional on $H_t(\underline{\theta}, \bar{\theta})$ is greater under $F$ than under $G$ that the probability distribution over $H_t(\underline{\theta}, \bar{\theta})$ under $F$ FOSD that under $G$.