SAPIENZA
UNIVERSITÀ DI ROMA

Discrete time models for bid-ask pricing under Dempster-Shafer uncertainty

PhD School of Economics
Department of Methods and Models for Economics, Territory and Finance
PhD programme in *Models for Economics and Finance* (XXXV cycle)

Andrea Cinfrignini
ID number 1895368

Advisor
Prof. Barbara Vantaggi

Co-Advisor
Prof. Davide Petturiti

Academic Year 2021/2022

Discrete time models for bid-ask pricing under Dempster-Shafer uncertainty
PhD thesis. Sapienza University of Rome

This thesis has been typeset by LaTeX and the Sapthesis class.

Author's email: andrea.cinfrignini@uniroma1.it

Abstract

As is well-known, real financial markets depart from simplifying hypotheses of classical no-arbitrage pricing theory. In particular, they show the presence of frictions in the form of bid-ask spread. For this reason, the aim of the thesis is to provide a model able to manage these situations, relying on a non-linear pricing rule defined as (discounted) *Choquet integral* with respect to a belief function.

Under the *partially resolving uncertainty* principle, we generalize the first fundamental theorem of asset pricing in the context of belief functions. Furthermore, we show that a generalized arbitrage-free *lower* pricing rule can be characterized as a (discounted) Choquet expectation with respect to an *equivalent inner approximating (one-step) Choquet martingale belief function.*

Then, we generalize the Choquet pricing rule dinamically: we characterize a *reference* belief function such that a multiplicative binomial process satisfies a suitable version of *time-homogeneity* and *Markov* properties and we derive the induced *conditional* Choquet expectation operator.

In a multi-period market with a risky asset admitting bid-ask spread, we assume that its lower price process is modeled by the proposed time-homogeneous Markov multiplicative binomial process. Here, we generalize the theorem of change of measure, proving the existence of an equivalent one-step Choquet martingale belief function. Then, we prove that the (discounted) lower price process of a European derivative is a *one-step Choquet martingale* and a *k-step Choquet super-martingale*, for $k \geq 2$.

Abstract

Come è ben noto, i mercati finanziari reali si discostano dalle ipotesi semplificative della teoria classica del pricing di non-arbitraggio. In particolare, essi mostrano la presenza di frizioni sotto forma di bid-ask spread. Per questa ragione, lo scopo della tesi è quello di fornire un modello in grado di gestire queste situazioni, facendo affidamento su un funzionale di prezzo non lineare definito come *integrale di Choquet* (scontato) rispetto ad una belief function.

Assumendo un contesto di *incertezza parzialmente risolvibile*, generalizziamo il primo teorema fondamentale dell'asset pricing, nel contesto delle belief functions. Inoltre, mostriamo che un funzionale di prezzo *inferiore* che soddisfa il no-arbitrage generalizzato può essere caratterizzato come valore atteso alla Choquet (scontato) rispetto a una *belief function equivalente e martingala di Choquet a un passo* ottenuta come approssimazione dall'interno.

Successivamente, generalizziamo il funzionale di prezzo di Choquet nel contesto dinamico: caratterizziamo una belief function di *riferimento* tale che un processo binomiale moltiplicativo soddisfi una versione (appropriata) delle proprietà di *stazionarietà* e *Markovianità*, e deriviamo il relativo valore atteso *condizionato* di Choquet.

In un mercato multiperiodale con un asset rischioso che ammette spread bid-ask, assumiamo che il suo processo di prezzo inferiore sia modellato dal processo binomiale moltiplicativo Markoviano e stazionario proposto. Qui, generalizziamo il teorema del cambiamento di misura, dimostrando l'esistenza di una belief function equivalente e martingala di Choquet a un passo. Quindi, dimostriamo che il processo di prezzo inferiore (scontato) di un derivato europeo è una *martingala* di Choquet a *un passo* e una *super-martingala* di Choquet a *k passi*, per $k \geq 2$.

Acknowledgments

Now that this life path is coming to an end, I can say that it has been the most challenging, also from a human standpoint, but also the most fulfilling experience.

I would like to start by thanking the reviewers who took the time to assess my work and provided me with helpful suggestions.

I want to express my special gratitude to Professor Davide Petturiti and Professor Barbara Vantaggi, who serve as my supervisors and role models, for their unending support.

Lastly, I desire to acknowledge those who have been and are in my life, or have been through it, because their influence and presence have shaped who I am.

Contents

Introduction

Most of the standard pricing models in discrete time suppose that uncertainty, both the *real-world* and the *risk-neutral*, is quantified by a unique probability measure. Starting from a real-world (or *natural*) probability measure that reflects the subjective uncertainty of agents, standard pricing theory assumes that the market is frictionless and competitive (Allingham, 1991). In this setting, the *first fundamental theorem of asset pricing* assures that a market is free of arbitrage opportunities if and only if there exists an equivalent martingale measure (Harrison and Kreps, 1979; Dybvig and Ross, 1989). Hence, the no-arbitrage assumption essentially materializes in the *linearity* and *time-consistency* of the dynamic pricing rule, which allows to define the price of any derivative which is not already traded in the market.

The *completeness* of the market guarantees that the equivalent martingale measure, and then the pricing rule, is *unique* (Pliska, 1997; Pascucci and Runggaldier, 2011). In a complete market, assuming the no-arbitrage condition, any derivative can be replicated by a *self-financing* strategy. The importance of the concept of completeness is due to the fact that it allows to price and hedge a derivative with a procedure that is preference-independent. This classical probabilistic results have been also justified through the game-theory setting (see, e.g., Shafer and Vovk, 2001, 2017, 2019).

As is well-known, the binomial model (Cox et al., 1979) is the simplest example of financial market that shows all features of no-arbitrage theory. Assuming that the no-arbitrage condition holds (under a suitable choice of parameters), the binomial model is complete, that is, there exists a *unique* equivalent martingale measure (also called *risk-neutral* measure) that enables the computation of no-arbitrage prices as a discounted expectation. As soon as we allow more than two possible values for the risky asset, obtaining a n-nomial model, completeness is lost. In this case, the no-arbitrage condition is equivalent to the existence of an infinite class of equivalent martingale measures and there are derivatives whose payoffs cannot be replicated by any self-financing strategy.

In the preceding example, incompleteness results from the fact that there are more sources of risk than (linearly independent) tradeable risky assets. Some techniques to face incompleteness look for a criterion to choose a particular $Q \in \mathcal{Q}$: for instance, Miyahara (1995) selects the probability measure that minimizes the relative entropy with respect to the natural probability. From the perspective of replication methods, the market is usually completed by adding the missing number of assets (see, e.g., Melnikov, 1999). Nevertheless, each *completion* uniquely defines an equivalent martingale measure in the set that arises from the incomplete market model. In turn, working with all possible completions of the market is equivalent to work with the class $\mathcal{Q}$ (Vasilev and Melnikov, 2021).

Since a *perfectly* replicating self-financing strategy may not exist, the *"best"* among imperfect ones can be chosen by means of approximations and algorithms. For instance, one can choose the strategy whose payoff is greater (less) than the derivative's payoff (*super(sub)-replicating strategy*); they result to be the no-arbitrage bounds for the non-replicable contract (Pascucci and Runggaldier, 2011). Other approximations can be achieved by minimizing, for instance, the *quadratic risk* (see, e.g., Bertsimas et al., 2001), or the *shortfall risk*, that penalizes only deviations in defect but it is less mathematically tractable (Cerný, 2009).

Besides incompleteness, real markets show other violations of the basic assumptions of classical no-arbitrage theory. Indeed, real markets are characterized by transaction costs, bid-ask spreads, portfolio constraints (see, e.g., Amihud and Mendelson, 1986, 1991), and situations of *ambiguity*, which occurs when we are not able to assign a specific probability measure that models uncertainty.

In this sense, the literature developed techniques to assess the impact of transaction costs on asset's equilibrium prices (Garman and Ohlson, 1980) in order to price European derivatives (Merton, 1990; Boyle and Vorst, 1992), or looked for a non-linear pricing rule defined through the envelopes of the class of equivalent martingale measures. In particular, the literature focused on a *lower (upper) pricing rule* expressed as a *(discounted) Choquet expectation* (Choquet, 1953) (see, e.g., Chateauneuf et al., 1996; Cerreia-Vioglio et al., 2015). This functional has the peculiarity that, when it is computed with respect to a particular non-additive measure (capacity) such as a *belief function* in Dempster-Shafer theory (Dempster, 1967; Shafer, 1976a), it is the lower envelope of expectations computed with respect to the dominating probabilities. In turn, working with the envelopes induced by the set of equivalent martingale measures is equivalent to work with the supremum and infimum self-financing portfolios computed, respectively, with respect to the set of all possible sub-replicating and super-replicating portfolios (Melnikov, 1999). Therefore, working with belief functions allows to introduce non-linearity departing the less from the classical no-arbitrage theory based on probability measures and to address dual representations in terms of strategies.

For a pricing theory to be accepted, a normative justification must be provided. For this, a generalization of the classical concept of arbitrage opportunity and the classical fundamental theorems of asset pricing are needed. In this vein, Carr et al. (2001) introduce two sets of probability measures (called *valuation* and *stress*) together with a definition of *no strictly acceptable opportunity* as the absence of an investment with zero price and expected value greater than zero, with respect to all probabilities in the sets. Under this condition, the pricing rule is a linear combination of the valuation probability measures. Their generalized second fundamental theorem of asset pricing, in turn, proves that if a market is *acceptable complete*, the linear combination is unique. In this way they enlarge the set of replicable derivatives. Also, the definitions of *ask* and *bid* prices are generalized, requiring that they are the smallest (greatest) value of the super(sub)-replicating portfolio comprising an acceptable loss of a given value.

This thesis inserts in this line of research. In detail, we present a research published in Cinfrignini et al. (2023) where we study a *one-period n-nomial market model*, that is already proved to be incomplete, by dealing with the whole class of equivalent martingale measures, each of them being consistent with the classical no-arbitrage assumption. We characterize the set of equivalent martingale measures and prove that the lower envelope of the closure of the class is a belief function. This suggests to directly work in the framework of belief functions which also allows to incorporate frictions in the market, in a natural way. Hence, our aim is to construct a *lower pricing rule* as a (discounted) lower expectation by means of the Choquet integral.

Anyhow, the closure of the set of equivalent martingale measures does not coincide with the credal set of the lower envelope, hence the (discounted) Choquet expectation with respect to the lower envelope is generally *lower* than the lower expectation computed with respect to the class of equivalent martingale measures.

After reformulating the one-period pricing problem over a finite state space in the framework of belief functions, we provide a generalization of the one-step no-Dutch book condition and of the one-step no-arbitrage condition with respect to a lower price assessment, based on the *partially resolving uncertainty* principle proposed by Jaffray (1989). Such principle allows to say that an agent may only acquire the information that an event occurs, without knowing which is the true state of the world. In turn, this translates in considering payoffs of portfolios on every event adopting a systematically pessimistic behaviour, that is always considering the minimum of random payoffs. This is in contrast to the usual *completely resolving uncertainty* assumption according to which the agent will always acquire which is the true state of the world and on which classical pricing theory is based.

We show that the generalized one-step no-Dutch book condition is necessary and sufficient for the existence of a belief function whose corresponding discounted Choquet expectation functional agrees with the lower price assessment, even though the positivity of the belief function cannot be guaranteed. The lack of positivity is an issue in the context of pricing since assets with non-negative and non-null payoff should have positive lower price. For this reason, we prove that the proposed generalized one-step no-arbitrage condition is equivalent to the existence of a *strictly positive* belief function whose corresponding discounted Choquet expectation functional agrees with the lower price assessment. The theorem we prove is the analogue of the *first fundamental theorem of asset pricing*, formulated in the context of belief functions and in a one-period setting. In particular, our result specializes results given in Chateauneuf et al. (1996); Cerreia-Vioglio et al. (2015), where the authors characterize an upper pricing rule that can be expressed as a discounted Choquet expectation with respect to a concave (or 2-alternating) capacity and require that a form of put-call parity relation has to hold. As already pointed out, working with belief functions in place of 2-monotone capacities (that are dual of 2-alternating capacities), allows to introduce non-linearity departing the less from the classical no-arbitrage setting.

Coming back to the original problem of specifying a lower pricing rule from the lower envelope arising in the n-nomial market model, the (discounted) Choquet expectation with respect to the lower envelope of the class of equivalent martingale measures does not satisfy the generalized no-Dutch book condition. To solve this issue, the idea is to ϵ*-contaminate* (Huber, 1981; Moral, 2018) a reference probability, suitably chosen in the set of equivalent martingale measures, with a belief function that is an *inner approximation* of the lower envelope and such that it satisfies the generalized no-arbitrage principle. The minimization procedure is achieved with respect to a suitable distance, similarly to what is done in Miranda et al. (2021, 2022); Montes et al. (2018, 2019). In this way we get an *equivalent inner approximating one-step Choquet martingale belief function* that is consistent with the lower price assessment of the market, it is equivalent to the (initial) lower envelope and it satisfies the generalized no-arbitrage principle. We conclude showing that, requiring that the equivalent Choquet martingale belief function is also consistent with the upper price assessment, it reduces to a probability measure.

Our analysis of the one-period n-nomial market model embodies frictions in the market by switching to the Dempster-Shafer theory. Hence, instead of defining a class of (martingale) probability measure and working with the corresponding envelopes, the next idea is to extend the one-period market model just defined to the multi-period case by taking the framework of belief functions as our natural environment to model uncertainty. Therefore, a necessary step is to define and characterize an *imprecise multiplicative binomial random process* in the multi-period setting with respect to a belief function such that it can be interpreted as a *lower* price process.

In order to define a multiplicative random process with respect to a belief function such that it is mathematically tractable, in analogy with the binomial process under a probability measure, we impose the *time-homogeneity* and *Markov* properties and we choose a suitable conditioning rule for belief functions that is the *product* conditioning rule (Suppes and Zanotti, 1977). Our study of binomial processes under Dempster-Shafer uncertainty started with the additive case (see Cinfrignini et al., 2022), that can be used to model log-returns. In the additive case, we defined a suitable version of the properties above, and proved that there exist some belief functions that satisfy just one of the desired properties (or that the property only holds one-step). Here, we characterize a *global* Markov and time-homogeneous multiplicative binomial random process (that we call *DS-multiplicative binomial process*) that differs from other proposals of imprecise Markov process that are in literature since they typically focus on local models rather than on a global one and they work with interval probabilities or capacities (in particular, we refer to Kast et al., 2014; Krak et al., 2019; Nendel, 2021; Škulj, 2016; T'Joens et al., 2021).

Since the usual Chapman-Kolmogorov equations do not hold in the Dempster-Shafer theory, we can characterize a multiplicative binomial process that meets the time-homogeneity and Markov properties by defining the structure of the *k-step transition belief functions*, that we additionally require to be interpretable and computationally tractable. We prove that a *global* belief function that meets all the desiderata exists and the transition belief functions are completely determined by the choice of only two parameters that can be interpreted as *"up"* and *"down"* one-step transition beliefs and such that when they sum up to one, the model reduces to the probabilistic one. This allows us to introduce a dynamic lower pricing rule expressed by a (discounted) *conditional* Choquet expectation operator of any function of a variable of the process with a closed form expression.

Thus, we consider a market composed by a frictionless risk-free bond and a non-dividend paying risky stock with frictions in the form of bid-ask spread, whose *lower* price process is modeled by the DS-multiplicative binomial process we characterized. In this market we prove a theorem that guarantees the existence of an *equivalent one-step Choquet martingale belief function* such that the (discounted) lower price process of the stock is a *one-step* Choquet *martingale* while, when more than one-step is considered, the lower price process is only a *global* Choquet *super-martingale*.

Next, considering the payoff of a European type derivative whose value depends only on the lower price of the stock, we propose a dynamic pricing rule that accounts for bid-ask spreads. The lower price process of the derivative is defined as a *one-step* (discounted) conditional Choquet expectation; the upper price process is defined one-step-wise through duality. For this pricing rule, the lower price process is shown to be always dominated by the upper price process and its discounted version turns out to be a Choquet super-martingale. Finally, we need to provide a normative justification of the proposed dynamic lower pricing rule by referring to a dynamic generalized no-arbitrage condition. Such condition, that is the extension to the multi-period setting of what we introduced in the one-period n-nomial model, based on the *partially resolving uncertainty* principle, justifies the choice of a parameter that is not consistent with the classical no-arbitrage condition but that is consistent with the existence of an equivalent one-step Choquet martingale belief function such that the generalized arbitrage opportunities are avoided.

The thesis is structured as follows.

Chapter 1 provides the mathematical foundations to introduce the models that follow, in particular the ambiguity and the Dempster-Shafer theory of belief functions.

Chapter 2 presents the classical no-arbitrage theory in discrete time models, focusing on the fundamental theorems and on the binomial and trinomial market models.

Chapter 3 reviews the main literature on incomplete markets, in particular the techniques to overcome the incompleteness in a market with transaction costs and bid-ask spreads, with non-linear pricing rules, together with the main literature on imprecise stochastic processes.

In Chapter 4 we present the work published in Cinfrignini et al. (2023), where the study of a n-nomial market model is developed, through the characterization of the envelopes of the set of equivalent martingale measures. We define a lower pricing rule by means of an equivalent inner approximating Choquet martingale belief function and we provide a generalization of the one-step no-arbitrage principle and of the relative theorems in the framework of belief functions.

Chapter 5 extends the one-period model of Chapter 4, studying a multiplicative binomial random process with respect to a belief function, asked to satisfy the time-homogeneity and Markov properties. Next, a dynamic Choquet pricing rule is proposed.

Finally, we gather conclusions and future perspectives.

Chapter 1

Uncertainty and ambiguity

In this Chapter we provide the mathematical foundations on which the following chapters are built. Then, we introduce the concept of *ambiguity* and we study non-additive measures, that generalize the classical probability measures, focusing on Dempster-Shafer theory of belief functions and on their conditioning rules proposed in literature.

1.1 Mathematical preliminaries

The set Ω is the non-empty set of all possible states of the world, called **sample space** (or certain event). In our framework, it is a *finite* set $\Omega = \{1, \dots, n\}$, with $n \in \mathbb{N}$, and we indicate by i a generic element (atom) in Ω and by A a generic subset of Ω, such that $A^C = \{i \in \Omega : i \notin A\}$ is the *complement* of the subset A. A collection of subsets of Ω, denoted by $\mathcal{F}$, is called **algebra** (or field) if it is closed under the formation of complements and finite unions, i.e. it satisfies the following conditions (Grimmett and Stirzaker, 2020):

(a) if $A, B \in \mathcal{F}$, then $A \cup B \in \mathcal{F}$ and $A \cap B \in \mathcal{F}$;

(b) if $A \in \mathcal{F}$, then $A^C \in \mathcal{F}$;

(c) $\emptyset \in \mathcal{F}$.

It follows that the *finite unions closure* holds:

$$A_1, \dots, A_n \in \mathcal{F} \Rightarrow \bigcup_{i=1}^{n} A_i \in \mathcal{F}. \tag{1.1}$$

A family of sets $B_1, B_2, \dots, B_n$ is called a **partition** of Ω if $B_i \cap B_j = \emptyset$ for $i \neq j$, and $\bigcup_{i=1}^{n} B_i = \Omega$.

If the state space Ω is *infinite*, the collection of events is required to be closed under operations of countable unions. A field that satisfies the *countable unions closure* is called **σ-algebra** (or σ-field):

$$A_1, A_2, \dots \in \mathcal{F} \Rightarrow A_1 \cup A_2 \cup \dots \in \mathcal{F}. \tag{1.2}$$

The smallest σ-field of Ω is $\mathcal{F} = \{\emptyset, \Omega\}$ while the σ-field that contains all subset of Ω is called **power set** and is denoted by $\mathcal{P}(\Omega) = 2^{\Omega}$. The pair $(\Omega, \mathcal{F})$ is called **measurable space**.

Since in the sequel we will restrict to a finite Ω, the notion of σ-algebra reduces to that of an algebra, thus the distinction becomes immaterial.

Uncertainty about the occurrences of events is generally modeled by a *probability measure.*

Definition 1.1 (Probability measure) *A probability measure P on the finite probability space $(\Omega, \mathcal{F})$ is a function $P : \mathcal{F} \to [0,1]$ such that:*

(1) $P(\emptyset) = 0$ and $P(\Omega) = 1$;

(2) for all $A, B \in \mathcal{F}$ with $A \cap B = \emptyset$

$$P(A \cup B) = P(A) + P(B).$$

The triple $(\Omega, \mathcal{F}, P)$ is called a probability space.

Note 1: *In case of an infinite Ω, condition (2) in Definition 1.1, that is referred to as finite additivity property, is usually replaced by the countable additivity property:*

(2') if $A_1, A_2, \ldots \in \mathcal{F}$ are pairwise disjoint (i.e. $A_i \cap A_j = \emptyset$ for all pairs i, j with $i \neq j$), then

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i).$$

Since in the sequel we deal with a finite Ω, countable additivity collapses into finite additivity and the distinction is actually immaterial.

An increasing sequence of σ-algebras $\{\mathcal{F}_t : t \in \mathcal{T}\}$, with $\mathcal{T} \subseteq \mathbb{R}$, such that $\mathcal{F}_t \subseteq \mathcal{F}_s \subseteq \mathcal{F}$, for all $t, s \in \mathcal{T}$, with $t < s$, is called a **filtration**. Each $\mathcal{F}_t$ represents the information available at time t and the filtration $\{\mathcal{F}_t\}_{t \in \mathcal{T}}$ is the information flow that increases with time. In what follows we assume that $\mathcal{T} = \{0, \ldots, T\}$ with $T \in \mathbb{N}$, and $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and $\mathcal{F}_T = \mathcal{P}(\Omega) = \mathcal{F}$. A probability space endowed with a filtration is called *filtered probability space* and it is denoted by $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in \mathcal{T}}, P)$.

Supposing that the "true" state of the world belongs to $B \subseteq \Omega$, with $B \neq \emptyset$, the **conditional probability** given B is customarily defined as

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \qquad \forall A \in \mathcal{F}, \tag{1.3}$$

provided that $P(B) > 0$. This definition of conditioning leaves $P(\cdot|B)$ undefined when $P(B) = 0$ and introduces the conditional measure as a "surrogate" of the unconditional measure. For this, starting from de Finetti (see, also, Dubins, 1975; Rényi, 1956) an axiomatic approach to conditional probability has been proposed, defining it as a two-place function $P : \mathcal{F} \times (\mathcal{F} \setminus \{\emptyset\}) \to [0,1]$ asked to satisfy:

(I) $P(A|B) = P(A \cap B|B)$, for every $A \in \mathcal{F}, B \in \mathcal{F} \setminus \{\emptyset\}$;

(II) $P(\cdot|B)$ is a (finitely additive) probability of $\mathcal{F}$, for every $B \in \mathcal{F} \setminus \{\emptyset\}$;

(III) $P(A \cap C|B) = P(A|B)P(C|A \cap B)$, for every $A, C \in \mathcal{F}$, $B, A \cap B \in \mathcal{F} \setminus \{\emptyset\}$.

A **random variable** on a probability space $(\Omega, \mathcal{F}, P)$ is a function $X : \Omega \to \mathbb{R}$ such that $\{i \in \Omega : X(i) \leq x\} \in \mathcal{F}$, for each $x \in \mathbb{R}$: it is said to be $\mathcal{F}$-measurable and we write $X \in \mathcal{F}$. In what follows we simply write $\{X \leq [\geq] x\} = \{i \in \Omega : X(i) \leq [\geq] x\}$. Let $A \in \mathcal{F}$ be an event, then $\mathbf{1}_A : \Omega \to \mathbb{R}$ denotes the indicator function of A.

A collection of random variables indexed by time $\{X_t : t \in \mathcal{T}\}$, with $\mathcal{T} = \{0, \dots, T\}$ and $T \in \mathbb{N}$, is called **random process** or **stochastic process**. Each X_t is a function $X_t : \Omega \to \mathcal{X}$, where $\mathcal{X}$ is called *state space*, such that for any $i \in \Omega$ there is a corresponding collection $\{X_t(i) : t \in \mathcal{T}\}$ of elements of $\mathcal{X}$ and it is called *realization* or *sample path* of X at i. A stochastic process is said to be *adapted* to the filtration $\{\mathcal{F}_t\}_{t \in \mathcal{T}}$ if each X_t is $\mathcal{F}_t$-measurable, i.e., $X_t \in \mathcal{F}_t$, for all $t \in \mathcal{T}$. There exists a filtration associated to each random process $\{\mathcal{F}_t^X\}$, that is generated by $\{X_t\}_{t \in \mathcal{T}}$, called *natural filtration* where $\mathcal{F}_t^X$ is the smallest σ-algebra that makes all random variables $\{X_s\}_{s \leq t}$ $\mathcal{F}_t$-measurable.

The **expectation** or **expected value** of a random variable X is defined to be

$$\mathbb{E}(X) = \int_\Omega X \,\mathrm{d}P = \int_0^{+\infty} P(X \geq t)\,\mathrm{d}t + \int_{-\infty}^0 [P(X \geq t) - 1]\,\mathrm{d}t. \tag{1.4}$$

Since Ω is finite, we actually have

$$\mathbb{E}(X) = \sum_{i \in \Omega} P(\{i\})X(i). \tag{1.5}$$

We specify that, in what follows, we will write $\mathbb{E}_P$ if we need to specify the probability measure P. For finite Ω, the expectation operator is a function $\mathbb{E} : \mathbb{R}^\Omega \to \mathbb{R}$ that satisfies the following properties:

(1) if $X \geq 0$, then $\mathbb{E}(X) \geq 0$;

(2) if $\alpha, \beta \in \mathbb{R}$, then $\mathbb{E}(\alpha X + \beta Y) = \alpha\mathbb{E}(X) + \beta\mathbb{E}(Y)$;

(3) if X and Y are independent, then $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$.

The **conditional expectation** of a random variable X on a finite filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in \mathcal{T}}, P)$, is defined as a random variable

$$\mathbb{E}(X|\mathcal{F}_t) = Y, \tag{1.6}$$

that is $\mathcal{F}_t$-measurable and that minimizes the quantity $\mathbb{E}[(X - Y)^2]$. Since Ω is finite, then

$$\mathbb{E}(X|\mathcal{F}_t)(i) = \int X \,\mathrm{d}P(\cdot|A), \tag{1.7}$$

for all $i \in A$ with A atom of $\mathcal{F}_t$.

The conditional expectation operator satisfies the **tower property**, or **law of iterated expectations**, that is defined, for $t \leq s$, as

$$\mathbb{E}\left[\mathbb{E}(X|\mathcal{F}_s)|\mathcal{F}_t\right] = \mathbb{E}(X|\mathcal{F}_t). \tag{1.8}$$

Definition 1.2 (Martingale) *Let $\{X_t : t \in \mathcal{T}\}$ be a random process on the probability space $(\Omega, \mathcal{F}, P)$ and let $\{\mathcal{F}_t : t \in \mathcal{T}\}$ be a filtration, then the sequence $(X_t, \mathcal{F}_t)$ is a martingale if, for all $s > t \geq 0$:*

(a) $\mathbb{E}(|X_t|) < \infty$*;*

(b) each $X_t \in \mathcal{F}_t$ *(the process is adapted to the filtration);*

(c) $\mathbb{E}(X_s|\mathcal{F}_t) = X_t$*.*

If $\mathcal{F}_t^X$ *is the natural filtration generated by the process, then obviously* X_t *is* $\mathcal{F}_t^X$*-measurable and the martingale property, for* $t \leq s$*, reduces to*

$$\mathbb{E}(X_s|\mathcal{F}_t^X) = \mathbb{E}(X_s|X_0, \dots, X_t) = X_t. \tag{1.9}$$

The process is called a ***super-martingale*** *if condition (c) holds with* $\leq$*; in turn, it is called a* ***sub-martingale*** *if condition (c) holds with* $\geq$*.*

Notice that, in case of a finite Ω, condition *(a)* is automatically satisfied.

Definition 1.3 (Markov process) *A discrete time random process* $\{X_t : t = 0, \dots, T\}$ *on a filtered probability space* $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t=0}^T, P)$*, adapted to the filtration* $\{\mathcal{F}_t\}_{t=0}^T$ *and where each* X_t *takes value in the state space* $\mathcal{X}$*, is said to be a Markov process if, for each* $t = 0, \dots, T-1$*, and* $1 \leq k \leq T-t$*, the distribution of* X_{t+k} *conditioned on* $\mathcal{F}_t$ *is the same as the distribution of* X_{t+k} *conditioned on* X_t*. This property can be written in the following equivalent ways:*

(a) (agreement of distributions) for every $i_0, \dots, i_{t+k} \in \mathcal{X}$ *we have*

$$\begin{aligned} P(X_{t+k} = i_{t+k}|\mathcal{F}_t) &= P(X_{t+k} = i_{t+k}|X_0 = i_0, \dots, X_t = i_t) \\ &= P(X_{t+k} = i_{t+k}|X_t = i_t) \\ &= \mathbb{E}(\mathbf{1}_i(X_{t+k})|X_t); \end{aligned}$$

(b) (agreement of expectations of all functions) for every function $h : \mathbb{R} \to \mathbb{R}$ *for which* $\mathbb{E}(|h(X_{t+k})|) < \infty$*, we have*

$$\mathbb{E}\left(h(X_{t+k})|\mathcal{F}_t\right) = \mathbb{E}\left(h(X_{t+k})|X_t\right).$$

In words, a Markovian random process is such that the *future is independent of the past, given the present value of the process*; the past values of the process can be ignored as long as we know the present state, indeed the Markov property is also called *memoryless property.*

The probability of crossing from a state $i \in \mathcal{X}$ at time t to state $j \in \mathcal{X}$ at time $t+1$ is called (one-step) **transition probability** and we denote it as

$$P(X_{t+1} = j|X_t = i) = p_{i,j}^{(t)}. \tag{1.10}$$

Definition 1.4 (Time-homogeneity) *A Markov process* $\{X_t\}$ *on a probability space* $(\Omega, \mathcal{F}, P)$ *is said to be (one-step) time-homogenous (or stationary) if transition probabilities do not depend on time*

$$P(X_{t+1} = j|X_t = i) = P(X_1 = j|X_0 = i) = p_{i,j}.$$

In a (one-step) stationary Markov process, transition probabilities are usually denoted as a matrix $\mathrm{P} = [p_{i,j}]$ that is squared and the number of rows is equal to the number of elements in $\mathcal{X}$. In order to define the k-step transition probabilities $P(X_{t+k} = j|X_t = i) = p_{i,j}^{(k)}$, i.e., the probability that a process in state i will be in state j after k steps, with $1 \leq k \leq T - t$, the **Chapman-Kolmogorov equations** are used (Ross, 2019). Denoting the initial probability with $P(X_0 = i) = \alpha_i$, the chain rule for conditional probabilities gives, for instance, for $T = 2$, the following result

$$\begin{gathered} P(X_0 = i_0, X_1 = i_1, X_2 = i_2) \\ = P(X_0 = i_0)P(X_1 = i_1|X_0 = i_0)P(X_2 = i_2|X_1 = i_1, X_0 = i_0) \\ = \alpha_{i_0} p_{i_0,i_1} p_{i_1,i_2}. \end{gathered}$$

In general, the following relation holds for $1 \leq t \leq T$ and $0 \leq k \leq T - t$,

$$P(X_{t+k} = j|X_t = i) = p_{i,i_{t+1}} p_{i_{t+1},i_{t+2}} \cdot \ldots \cdot p_{i_{t+k-1},j}. \tag{1.11}$$

It follows that the k-th order transition probabilities are given by

$$p_{i,j}^{(k)} = P(X_{t+k} = j|X_t = i) = \sum_{i_1,\ldots,i_{k-1}} p_{i,i_1} \cdot \ldots \cdot p_{i_{k-1},j}. \tag{1.12}$$

1.1.1 Theorems of the alternative

Theorems of the alternative show that a linear system can be associated with a dual system of constraints such that one system is feasible if and only if the other one is infeasible; in other words, they state that a condition is true if and only if the other is false. We provide a summary of the most useful theorems since they will be used in proofs and applications in the following chapters, referring to Ben-Israel (2001).

Let us consider two matrices $A \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{m\times n}$ with $rank(A) = m$, and vectors $\boldsymbol{x} \in \mathbb{R}^n, \boldsymbol{b} \in \mathbb{R}^m$, $\boldsymbol{c} \in \mathbb{R}^m$. Note that $\mathbf{0}$ denotes a vector of zeros in $\mathbb{R}^n$. Linear problems can appear in different forms (called *primal*) in which the following are the typical (Ben-Israel, 2001):

$$A\boldsymbol{x} \leq \boldsymbol{b}; \tag{[a]}$$

$$A\boldsymbol{x} = \boldsymbol{b},\ \boldsymbol{x} \geq \mathbf{0}; \tag{[b]}$$

$$A\boldsymbol{x} \leq \boldsymbol{b},\ B\boldsymbol{x} < \boldsymbol{c}; \tag{[c]}$$

$$A\boldsymbol{x} = \mathbf{0},\ \boldsymbol{x} \gneqq \mathbf{0}; \tag{[d]}$$

$$A\boldsymbol{x} = \mathbf{0},\ \boldsymbol{x} > \mathbf{0}; \tag{[e]}$$

$$A\boldsymbol{x} > \mathbf{0},\ B\boldsymbol{x} \geq \mathbf{0},\ C\boldsymbol{x} = \mathbf{0}. \tag{[f]}$$

First of all Ben-Israel (2001) stresses that all primal representations reduces to ([c]). In fact, ([a]) and ([b]) are special case of ([c]). Moreover, ([d]) and ([e]) are special case of ([f]). In turn, ([f]) is a special case of ([c]) with $\boldsymbol{b} = \mathbf{0}$ and $\boldsymbol{c} = \mathbf{0}$. Hence, each linear problem can be written as ([c]).

Theorem 1.1 (Motzkin's Transposition Theorem)
Given matrices $A, B \in \mathbb{R}^{m\times n}$ and vectors $\boldsymbol{b}, \boldsymbol{c} \in \mathbb{R}^m$, then the following are equivalent:

(1) the system ([c]) $A\boldsymbol{x} \leq \boldsymbol{b},\ B\boldsymbol{x} < \boldsymbol{c}$ has a solution $\boldsymbol{x} \in \mathbb{R}^n$;

(2) for all vectors $\boldsymbol{y}, \boldsymbol{z} \in \mathbb{R}^m$ with $\boldsymbol{y} \geq \boldsymbol{0}$ and $\boldsymbol{z} \geq \boldsymbol{0}$,
$A^T\boldsymbol{y} + B^T\boldsymbol{z} = \boldsymbol{0},\ \boldsymbol{z} = \boldsymbol{0} \Rightarrow \boldsymbol{b}^T\boldsymbol{y} + \boldsymbol{c}^T\boldsymbol{z} \geq 0$ and
$A^T\boldsymbol{y} + B^T\boldsymbol{z} = \boldsymbol{0},\ \boldsymbol{z} \neq \boldsymbol{0} \Rightarrow \boldsymbol{b}^T\boldsymbol{y} + \boldsymbol{c}^T\boldsymbol{z} > 0$.

The Farkas's Theorems are special cases of the Motzkin's Theorem, in particular for solving systems ([a]) and ([b]).

Theorem 1.2 (Farkas's Theorem ([a]))
Given a matrix $A \in \mathbb{R}^{m\times n}$ and a vector $\boldsymbol{b} \in \mathbb{R}^m$, the following are equivalent

(a) the system ([a]) $A\boldsymbol{x} \leq \boldsymbol{b}$ has a solution $\boldsymbol{x} \in \mathbb{R}^n$;

(b) for all $\boldsymbol{y} \in \mathbb{R}^m$,
$A^T\boldsymbol{y} = \boldsymbol{0},\ \boldsymbol{y} \geq \boldsymbol{0} \Rightarrow \boldsymbol{b}^T\boldsymbol{y} \geq 0$.

Theorem 1.3 (Farkas's Theorem ([b]))
Given a matrix $A \in \mathbb{R}^{m\times n}$ and a vector $\boldsymbol{b} \in \mathbb{R}^m$, the following are equivalent

(i) the system ([b]) $A\boldsymbol{x} = \boldsymbol{b},\ \boldsymbol{x} \geq \boldsymbol{0}$ has a solution $\boldsymbol{x} \in \mathbb{R}^n$;

(ii) for all $\boldsymbol{y} \in \mathbb{R}^m$,
$A^T\boldsymbol{y} \geq \boldsymbol{0} \Rightarrow \boldsymbol{b}^T\boldsymbol{y} \geq 0$.

1.2 Ambiguity and Dempster-Shafer theory

In real world, an agent is usually not able to assign a specific probability measure to the events of interest, while she/he often needs to consider a set of probability measures or an imprecise probability measure. In the literature this situation is called *ambiguity* (Epstein and Schneider, 2010, 2007; Gilboa and Marinacci, 2016; Epstein and Schneider, 2008; Pennesi, 2018; Klibanoff et al., 2009; Gilboa and Schmeidler, 1989). The seminal paper of Ellsberg (1961) shows that, in a setting in which agents have a set of probability measures resulting from imperfect information (that represents the prototype of ambiguity), the agents' behaviour cannot be justified in terms of expected utility.

On the contrary, this violation can be supported by the presence of *ambiguity adverse* agents who prefer a situation of known probabilities rather than unknown.

The following motivating example (see also Coletti et al., 2015), based on the Ellsberg paradox (Ellsberg, 1961) and originally set up in a decision theoretic framework, shows a situation of partial knowledge that leads to a set of probabilities, with respect to which we compute the envelopes.

Example 1.1 *Suppose to have an urn from which we draw a ball. The urn contains $\frac{1}{3}$ of white balls (w) and the remaining balls are black (b) and red (r) in a proportion that is completely unknown to us. We denote by $\Omega = \{w, b, r\}$ and with $\mathcal{P}(\Omega) = \{\emptyset, \{w\}, \{b\}, \{r\}, \{w, b\}, \{w, r\}, \{b, r\}, \Omega\}$. The composition of the urn leads to the following probabilities:*

$$P_\lambda(\{w\}) = \frac{1}{3}, \quad P_\lambda(\{b\}) = \lambda, \quad P_\lambda(\{r\}) = \frac{2}{3} - \lambda,$$

with $\lambda \in \left[0, \frac{2}{3}\right]$. *This means that, by varying* λ, *we have a class of infinitely many probability measures*

$$\mathcal{P} = \left\{P_\lambda : \lambda \in \left[0, \frac{2}{3}\right]\right\}.$$

Since we cannot work with an infinite number of probability measures, we compute the lower and the upper envelope of the class, respectively denoted by $\underline{P} = \min \mathcal{P}$ *and* $\overline{P} = \max \mathcal{P}$ *where minimum and maximum are computed pointwise on* $\mathcal{P}(\Omega)$. *Results are reported below*

$\mathcal{P}(\Omega)$	$\emptyset$	$\{w\}$	$\{b\}$	$\{r\}$	$\{w,b\}$	$\{w,r\}$	$\{b,r\}$	Ω
$P_{\lambda=0}$	0	$\frac{1}{3}$	0	$\frac{2}{3}$	$\frac{1}{3}$	1	$\frac{2}{3}$	1
$P_{\lambda=2/3}$	0	$\frac{1}{3}$	$\frac{2}{3}$	0	1	$\frac{1}{3}$	$\frac{2}{3}$	1
$\underline{P}$	0	$\frac{1}{3}$	0	0	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{2}{3}$	1
$\overline{P}$	0	$\frac{1}{3}$	$\frac{2}{3}$	$\frac{2}{3}$	1	1	$\frac{2}{3}$	1

It is easily verified that the lower envelope $\underline{P}$ *and the upper envelope* $\overline{P}$ *are not probability measures since they do not satisfy (2) in Definition 1.1.*

♦

Instead of choosing one specific probability measure from the class (that would require to specify a criterion of choice such as using the agent's subjective probability, as the Bayesian paradigm prescribes), our approach is to work with the entire class of probability measures through its envelopes. The envelopes of a set of probability measures are typically not probabilities, though. Thus, the probabilistic framework presented in the preceding section does not apply in this situation, necessitating a generalisation because conditioning and expectation refer to additive probability.

Therefore, we need to introduce non-additive measures such that they reduce to probability if the additive property (*(2)* of Definition 1.1) is satisfied.

Belief functions, proposed by Dempster and Shafer (Dempster, 1967; Shafer, 1976a)[1], are one example of non-additive measures that require to assign a degree of belief to every subset of the sample space, rather than focusing just to atoms, as a probability measure does.

1.2.1 Non-additive measures

Let $(\Omega, \mathcal{F})$ be a finite measurable space with $\mathcal{F} = \mathcal{P}(\Omega)$. A subset $A \in \mathcal{F}$ is an event such that a function $\nu(A)$ quantifies the amount of uncertainty that the event A realizes, in other words, that the true state of nature lies in A. In the starting section we considered that ν was a probability measure P. However, the probabilistic information can be incomplete or imprecise, then a generalization of the probabilistic setting has to be provided.

Definition 1.5 (Non-additive measure) *A function* $\nu : \mathcal{F} \to \mathbb{R}$ *is called a non-additive measure or a capacity if it is:*

[1] In some cases we will shorten "Dempster-Shafer" as "DS".

(i) (normalized) $\nu(\emptyset) = 0,$ *and* $\nu(\Omega) = 1$*;*

(ii) (monotone) $\nu(A) \leq \nu(B)$ *for all* $A, B \in \mathcal{F}$*, with* $A \subseteq B$*.*

The capacity ν *is called:*

(a) ***2-monotone*** *(or convex, supermodular capacity) if, for every* $A, B \in \mathcal{F}$*,*

$$\nu(A \cup B) \geq \nu(A) + \nu(B) - \nu(A \cap B);$$

(b) ***k-monotone*** *if, for every* $A_1, \ldots, A_k \in \mathcal{F}$*,*

$$\nu\left(\bigcup_{i=1}^{k} A_i\right) \geq \sum_{\emptyset \neq I \subseteq \{1,\ldots,k\}} (-1)^{|I|+1} \nu\left(\bigcap_{i \in I} A_i\right);$$

(c) ***completely monotone*** *or a* ***belief function****, denoted by* Bel*, if, for every* $A_1, \ldots, A_k \in \mathcal{F}$ *with* $k \geq 2$*,*

$$\nu\left(\bigcup_{i=1}^{k} A_i\right) \geq \sum_{\emptyset \neq I \subseteq \{1,\ldots,k\}} (-1)^{|I|+1} \nu\left(\bigcap_{i \in I} A_i\right);$$

(d) ***necessity measure*** *if, for every* $A, B \in \mathcal{F}$*,*

$$\nu(A \cap B) = \min\{\nu(A), \nu(B)\};$$

(e) (coherent) ***lower probability*** *if there exists a set* $\mathcal{P}$ *of probability measures on* $\mathcal{F}$ *such that, for every* $A \in \mathcal{F}$*,*

$$\nu(A) = \inf_{P \in \mathcal{P}} P(A).$$

(f) ***probability measure*** *if, for every* $A, B \in \mathcal{F}$ *such that* $A \cap B = \emptyset$*,*

$$\nu(A \cup B) = \nu(A) + \nu(B).$$

A k-monotone capacity is also k'-monotone for $2 \leq k' \leq k$ and a k-monotone capacity, for each k, is a (coherent) lower probability. Necessity measures (denoted as N) are particular belief functions and belief functions are particular (coherent) lower probabilities. On the other hand, lower order monotonicities do not imply higher order monotonicities.

We denote by $\mathbf{V}(\Omega, \mathcal{F})$, $\mathbf{B}(\Omega, \mathcal{F})$ and $\mathbf{P}(\Omega, \mathcal{F})$, respectively, the set of all capacities, belief functions and probability measures on $(\Omega, \mathcal{F})$, and we point out that $\mathbf{P}(\Omega, \mathcal{F}) \subset \mathbf{B}(\Omega, \mathcal{F}) \subset \mathbf{V}(\Omega, \mathcal{F})$.

For every capacity ν there exists a function called *conjugate function* or *dual capacity*, denoted by $\overline{\nu} : \mathcal{F} \to \mathbb{R}$, such that, for all $A \in \mathcal{F}$, it is

$$\overline{\nu}(A) = 1 - \nu(A^C). \tag{1.13}$$

The dual of a lower probability is called *upper probability*; the dual of a k-monotone capacity is said *k-alternating capacity*; the dual of a belief function is said *plausibility* *function* (denoted as Pl); the dual of a necessity measure is said *possibility measure* (denoted as Π); the dual of a probability is itself: this property is called *self-duality*.

For every (coherent) lower probability there exists a non-empty set of dominating probability measures called **core** (or *credal set*) (Gilboa and Schmeidler, 1994; Walley, 1991; Denneberg, 1994)

$$\mathrm{core}(\nu) = \{P \in \mathbf{P}(\Omega, \mathcal{F}) \ : \ P(A) \geq \nu(A), \ \forall A \in \mathcal{F}\}\,, \tag{1.14}$$

such that $\nu = \min \mathrm{core}(\nu)$.

If ν is a (coherent) lower probability determined by the class of probabilities $\mathcal{P}$ on $\mathcal{F}$, then we get a *lower expectation operator* by setting, for every random variable $X \in \mathbb{R}^\Omega$

$$\underline{\mathbb{E}}_{\mathcal{P}}(X) = \inf_{P \in \mathcal{P}} \mathbb{E}_P(X). \tag{1.15}$$

In particular, by taking the closure of $\mathcal{P}$ with respect to the product topology $\mathrm{cl}(\mathcal{P})$, infima are attained, that is

$$\underline{\mathbb{E}}_{\mathcal{P}}(X) = \min_{P \in \mathrm{cl}(\mathcal{P})} \mathbb{E}_P(X). \tag{1.16}$$

In general, since $\mathrm{cl}(\mathcal{P}) \subseteq \mathrm{core}(\nu)$, then

$$\underline{\mathbb{E}}_{\nu}(X) = \min_{P \in \mathrm{core}(\nu)} \mathbb{E}_P(X) \leq \min_{P \in \mathrm{cl}(\mathcal{P})} \mathbb{E}_P(X) = \underline{\mathbb{E}}_{\mathcal{P}}(X). \tag{1.17}$$

If ν has non-empty core but it is not a lower probability, we can still introduce a lower expectation operator said *natural extension*, in the jargon of Walley,

$$\underline{\mathbb{E}}_{\nu}^{N}(X) = \min_{P \in \mathrm{core}(\nu)} \mathbb{E}_P(X). \tag{1.18}$$

Notice that $\underline{\mathbb{E}}_{\nu}^{N}(\mathbf{1}_A) \geq \nu(A)$, for all $A \in \mathcal{F}$.

Working with the dual function $\overline{\nu}$, there exists a non-empty set of *dominated* probability measures, called $\mathrm{anticore}(\overline{\nu})$, such that

$$\mathrm{anticore}(\overline{\nu}) = \{P \in \mathbf{P}(\Omega, \mathcal{F}) \ : \ P(A) \leq \overline{\nu}(A), \ \forall A \in \mathcal{F}\}. \tag{1.19}$$

If ν (and then $\overline{\nu}$) are 2-monotone (2-alternating) capacities, then $\mathrm{core}(\nu) = \mathrm{anticore}(\overline{\nu})$. To have a lighter notation, in what follows we will often refer to core also in case of a 2-alternating capacity, always intending the core of its dual.

If ν is a 2-monotone capacity, the set of extreme points of the core is

$$\mathrm{ext}(\mathrm{core}(\nu)) = \{\nu^\sigma \ : \ \sigma \in \Sigma\},$$

where Σ is the set of all permutations of indices $\{1, \ldots, n\}$, and

$$\nu^\sigma = (\nu\left(\sigma(1)\right), \ldots, \nu\left(\sigma(n)\right))\,, \tag{1.20}$$

with $\nu(\sigma(i)) = \nu(\{\sigma(1), \ldots, \sigma(i)\}) - \nu(\{\sigma(1), \ldots, \sigma(i-1)\})$.

We stress that ν^σ gives rise to a probability measure (Destercke and Dubois, 2014).

As already mentioned, a coherent lower probability is the lower envelope of its core and the same property holds if the lower probability reduces to a belief function, i.e., for all $A \in \mathcal{F}$ holds that

$$\underline{P}(A) = \min_{P \in \text{core}(\underline{P})} P(A), \tag{1.21}$$

$$Bel(A) = \min_{P \in \text{core}(Bel)} P(A). \tag{1.22}$$

Each capacity can be expressed in terms of another function called *Möbius inverse*.

Definition 1.6 (Möbius inverse) *For any capacity $\nu : \mathcal{F} \to \mathbb{R}$ there is a one-to-one correspondence with another function called Möbius inverse or mass function $\mu : \mathcal{F} \to \mathbb{R}$, where*

$$\mu(A) = \sum_{B \subseteq A} (-1)^{|A \setminus B|} \nu(B), \qquad \nu(A) = \sum_{B \subseteq A} \mu(B). \tag{1.23}$$

It implies that $\mu(\emptyset) = 0$ and $\sum_{A \in \mathcal{F}} \mu(A) = 1$. Events $A \in \mathcal{F}$ such that $\mu(A) > 0$ are called *focal elements* and $\mu(i) = \nu(i) \geq 0$, $\forall i \in \Omega$.

Properties of the capacity ν reflect into properties of its Möbius inverse μ.

Proposition 1.1 (Chateauneuf and Jaffray, 1989)
Given a function $\nu : \mathcal{F} \to \mathbb{R}$ and its Möbius inverse $\mu : \mathcal{F} \to \mathbb{R}$, then

(a) ν is a **capacity** if and only if

- *(a.1)* $\mu(\emptyset) = 0$;
- *(a.2)* $\sum_{B \in \mathcal{F}} \mu(B) = 1$;
- *(a.3)* $\sum_{\{i\} \subseteq B \subseteq A} \mu(B) \geq 0$, for all $A \in \mathcal{F}$ and for all $i \in A$;

(b) ν is a **2-monotone** capacity if and only if condition *(a)* holds and, $\forall A \in \mathcal{F}$ and $\{i, j\} \subseteq A$ with $i \neq j$,

$$\sum_{\{i,j\} \subseteq B \subseteq A} \mu(B) \geq 0;$$

(c) ν is a **belief function** if and only if condition *(a)* holds and μ is non-negative: $\mu(A) \geq 0$, for all $A \in \mathcal{F}$;

(d) ν is a **necessity measure** if and only if condition *(a)* holds and the set of its focal elements is totally ordered by the inclusion relation;

(e) ν is a **probability measure** if and only if condition *(a)* holds, μ is non-negative and it can be positive only on singletons.

We note that, dealing with belief functions, the Möbius inverse is also called *basic probability assignment* and $\mu(A)$ can be interpreted as the evidence that the true value of a random variable is in $A \in \mathcal{F}$ (Shafer, 1976a).

Example 1.1 showed that non-additive measures can be achieved as envelopes of a set of known probabilities; in particular, the lower envelope $\underline{P}$ in the example is a belief function. In turn, a belief function, through its Möbius inverse, can be directly assigned by an agent to events, based on the evidence in favour or not, such that this can easily point out the degree of ignorance. In the following example, inspired by Boivin (2019), we show this approach.

Example 1.2 *Let's say that we are interested in the weather in Rome and, more specifically, we want to know how many days there will be sunshine or not in a month, hence our states of the world can be only "sun" and "not sun", respectively denoted by* $\Omega = \{\{s\}, \{\overline{s}\}\}$. *Sadly, there are no historical data on the number of sunny days but there are on the number of rainy days, that are* 45% *of the month's days. Since we are interested about the sun, if it rains, there is not sun, hence we assign a mass to the event "not sun"*

$$\mu(\{\overline{s}\}) = 0.45.$$

Conversely, event "no rain" does not mean "sun", but, for instance, it can be cloudy. Hence "no rain" is compatible with the event {"sun", "not sun"}. Then we assign the remaining mass to

$$\mu(\{s, \overline{s}\}) = \mu(\Omega) = 1 - 0.45 = 0.55.$$

Let us compute the associated capacity, that results to be a belief function since μ *satisfies condition (c) of Proposition 1.1*

	$\emptyset$	$\{s\}$	$\{\overline{s}\}$	Ω
μ	0	0	0.45	0.55
Bel	0	0	0.45	1

Hence, $Bel(\{s\}) = 0$ *is the evidential support in favour of "sun" and* $Bel(\{\overline{s}\}) = 0.45$ *is the evidential support in favour of "not sun". We can easily verify that* Bel *is not a probability measure since it does not satisfy the additivity property*

$$Bel(s \cup \overline{s}) = Bel(\{s, \overline{s}\}) = 1 \geq Bel(\{s\}) + Bel(\{\overline{s}\}) = 0.45.$$

♦

The above setting can be equivalently expressed in terms of *coherent lower expectation (prevision).*

Definition 1.7 ((coherent) lower expectation) *A coherent lower expectation (prevision)* $\underline{\mathbb{E}} : \mathbb{R}^\Omega \to \mathbb{R}$, *is characterized by the following axioms:*

(E.1) $\inf_{i \in \Omega} X(i) \leq \underline{\mathbb{E}}(X) \leq \sup_{i \in \Omega} X(i)$;

(E.2) (positive homogeneity) $\underline{\mathbb{E}}(\alpha X) = \alpha \underline{\mathbb{E}}(X)$, *with* $\alpha > 0$;

(E.3) (super-linearity) $\underline{\mathbb{E}}(X + Y) \geq \underline{\mathbb{E}}(X) + \underline{\mathbb{E}}(Y)$.

The lower probability of an event $A \in \mathcal{F}$ is defined as the lower expectation of its indicator $\nu(A) = \underline{\mathbb{E}}(\mathbf{1}_A)$.

In analogy with coherent lower probability, any coherent lower expectation $\underline{\mathbb{E}}$ has a *conjugate (coherent) upper expectation* defined, for all $X \in \mathbb{R}^\Omega$, by

$$\overline{\mathbb{E}}(X) = -\underline{\mathbb{E}}(-X). \tag{1.24}$$

Moreover, any coherent lower expectation induces a closed and convex set of *compatible* expectations

$$\mathcal{M}\left(\underline{\mathbb{E}}\right) = \left\{\mathbb{E} \text{ expectation } : \forall X \in \mathbb{R}^{\Omega}, \underline{\mathbb{E}}(X) \leq \mathbb{E}(X) \leq \overline{\mathbb{E}}(X)\right\}, \tag{1.25}$$

such that $\underline{\mathbb{E}}(X) = \min_{\mathbb{E}\in\mathcal{M}(\underline{\mathbb{E}})} \mathbb{E}(X)$, and $\overline{\mathbb{E}}(X) = \max_{\mathbb{E}\in\mathcal{M}(\underline{\mathbb{E}})} \mathbb{E}(X)$.

1.2.2 Choquet integral

Within the framework of probability measures, equation (1.4) showed how to compute the expected value of a random variable. Once that additivity is lost, we must design a functional that yields the same meaning of the expected value and that reduces to it when ν is additive.

The answer is not unique and, in what follows, we consider the Choquet integral (Choquet, 1953) that can be computed with respect to any capacity ν.

Let us consider a measurable finite space $(\Omega, \mathcal{F})$. We recall that $\mathbb{R}^{\Omega}$ denotes the set of all real-valued random variables on Ω.

Definition 1.8 (Choquet integral) *Given a capacity ν and a random variable $X \in \mathbb{R}^{\Omega}$, the Choquet integral of X with respect to ν, denoted by $\mathbb{C}_{\nu}(X)$, is defined as*

$$\mathbb{C}_{\nu}(X) = \oint_{\Omega} X \,\mathrm{d}\nu = \int_{0}^{\infty} \nu(X \geq t)\,\mathrm{d}t + \int_{-\infty}^{0} \left[\nu(X \geq t) - 1\right]\,\mathrm{d}t.$$

Since we assume a finite Ω and $\mathcal{F} = \mathcal{P}(\Omega)$, the Choquet integral can be computed as

$$\mathbb{C}_{\nu}(X) = \sum_{i=1}^{n} \left[X(\sigma(i)) - X(\sigma(i+1))\right] \nu(E_i^{\sigma}), \tag{1.26}$$

where σ is a permutation of Ω such that $X(\sigma(1)) \geq \cdots \geq X(\sigma(n))$, $X(\sigma(n+1)) = 0$, and $E_i^{\sigma} = \{\sigma(1), \ldots, \sigma(i)\}$, for $i = 1, \ldots, n$.

If ν reduces to a probability measure P, the Choquet integral reduces to the expected value of X

$$\mathbb{C}_{P}(X) = \oint_{\Omega} X \,\mathrm{d}P = \int_{0}^{\infty} P(X \geq t)\,\mathrm{d}t + \int_{-\infty}^{0} \left[P(X \geq t) - 1\right]\,\mathrm{d}t = \mathbb{E}_{P}(X).$$

Moreover, for every capacity $\nu \in \mathbf{V}(\Omega, \mathcal{F})$, the Choquet integral can be computed through the corresponding Möbius inverse of ν.

Theorem 1.4 (Gilboa and Schmeidler, 1994)
For every $\nu \in \mathbf{V}(\Omega, \mathcal{F})$ with corresponding μ, and $X \in \mathbb{R}^{\Omega}$, it holds that

$$\mathbb{C}_{\nu}(X) = \sum_{B \in \mathcal{F}\setminus\{\emptyset\}} \mu(B) \left(\min_{i \in B} X(i)\right).$$

Now we mention some properties of the Choquet integral that will be helpful in the next sections. Let ν be any capacity in $\mathbf{V}(\Omega, \mathcal{F})$, and X a random variable in $\mathbb{R}^{\Omega}$, then:

(Ch.1) for all $A \subseteq \Omega$, we have that $\mathbb{C}_\nu(\mathbf{1}_A) = \nu(A)$, where $\mathbf{1}_A$ is the indicator function of A;

(Ch.2) for any other capacity $\varphi \in \mathbf{V}(\Omega, \mathcal{F})$, $X \in \mathbb{R}^\Omega$ and $\alpha, \beta \in \mathbb{R}$, it holds that

$$\mathbb{C}_{\alpha\nu+\beta\varphi}(X) = \alpha\mathbb{C}_\nu(X) + \beta\mathbb{C}_\varphi(X);$$

(Ch.3) (*non-negative homogeneity*) for all $\alpha \geq 0$ and $X \in \mathbb{R}^\Omega$, it holds that

$$\mathbb{C}_\nu(\alpha X) = \alpha\mathbb{C}_\nu(X);$$

(Ch.4) (*constant additivity*) for all $\alpha \in \mathbb{R}$, it holds that

$$\mathbb{C}_\nu(\alpha + X) = \alpha + \mathbb{C}_\nu(X);$$

(Ch.5) (*monotonicity*) for all $X, Y \in \mathbb{R}^\Omega$ such that $X \leq Y$, it holds that

$$\mathbb{C}_\nu(X) \leq \mathbb{C}_\nu(Y);$$

(Ch.6) (*comonotone additivity*) if $X, Y \in \mathbb{R}^\Omega$ are *comonotone*, meaning that for all $i, j \in \Omega$ $[X(i) - X(j)][Y(i) - Y(j)] \geq 0$, it holds that

$$\mathbb{C}_\nu(X + Y) = \mathbb{C}_\nu(X) + \mathbb{C}_\nu(Y);$$

(Ch.7) (*asymmetry*) if $\overline{\nu}$ is the dual capacity of ν, it holds that

$$\mathbb{C}_\nu(-X) = -\mathbb{C}_{\overline{\nu}}(X);$$

(Ch.8) if ν reduces to a 2-monotone capacity, for all $X, Y \in \mathbb{R}^\Omega$, the Choquet integral is a coherent lower expectation as given in Definition 1.7, i.e. it is *super-additive*

$$\mathbb{C}_\nu(X) + \mathbb{C}_\nu(Y) \leq \mathbb{C}_\nu(X + Y),$$

and, combined with non-negative homogeneity, it satisfies the *super-linearity* property[2]. Moreover, the Choquet integral computed with respect to (at least) a 2-monotone capacity is the lower expectation with respect to the core of the capacity (we point out that this property continues to hold if ν reduces to a belief function) (Gilboa and Schmeidler, 1994)

$$\mathbb{C}_\nu(X) = \min_{P \in \mathrm{core}(\nu)} \sum_{i \in \Omega} P(\{i\})X(i) = \min_{P \in \mathrm{core}(\nu)} \mathbb{E}_P(X) = \underline{\mathbb{E}}(X). \tag{1.27}$$

It follows that the Choquet integral coincides with the natural extension (1.18).[3]

[2]We stress that for a 2-alternating (concave) capacity, the Choquet integral is *sub-additive* and, combined with non-negative homogeneity, it is *sub-linear*.

[3]Notice that, if ν is a (coherent) lower probability not 2-monotone, then $\mathrm{core}(\nu) \neq \emptyset$ but there exists at least a random variable X such that the Choquet integral is smaller than the natural extension: $\exists X \in \mathbb{R}^\Omega \ : \ \mathbb{C}(X) < \underline{\mathbb{E}}^N(X)$.

In turn, given a family of probabilities $\mathcal{P}$, if its lower envelope $\nu = \min \mathrm{cl}(\mathcal{P})$ is a 2-monotone (or completely monotone) capacity, but $\mathrm{cl}(\mathcal{P})$ is strictly contained in $\mathrm{core}(\nu)$, then the equality (1.27) may fail to hold. Indeed, computing the Choquet integral with respect to ν we are actually computing the lower expectation functional on $\mathbb{R}^\Omega$ determined by $\mathrm{core}(\nu)$, which is dominated by the lower expectation functional on $\mathbb{R}^\Omega$ determined by $\mathcal{P}$ (see, e.g., de Cooman et al., 2008). This latter functional is 2-monotone (completely monotone) on indicators of events but may fail 2-monotonicity (complete monotonicity) on the whole $\mathbb{R}^\Omega$. This means that for some $X \in \mathbb{R}^\Omega$ we can have $\mathbb{C}_\nu(X) < \min_{P \in \mathrm{cl}(\mathcal{P})} \mathbb{E}_P(X)$.

(Ch.9) (*complete monotonicity*) if ν reduces to a belief function Bel, for all $k \geq 2$ and all $X_1, \ldots, X_k \in \mathbb{R}^\Omega$, the Choquet integral is *completely monotone*

$$\mathbb{C}_{Bel}\left(\bigvee_{i=1}^{k} X_i\right) \geq \sum_{\emptyset \neq I \subseteq \{1,\ldots,k\}} (-1)^{|I|+1} \mathbb{C}_{Bel}\left(\bigwedge_{i \in I} X_i\right). \tag{1.28}$$

1.2.3 Conditioning

Working with generalized measures requires also an adequate generalization of the conditioning rule in the non-additive measure's environment since it is fundamental to develop dynamic pricing models. Literature proposed several definitions of conditional capacities which continue to hold when we deal with a belief function. The following conditioning rules are proposed in terms of belief and plausibility functions even if they can be stated more generally in terms of capacities (Denneberg, 1994; Grabish, 2016): this is because we will mainly work in the Dempster-Shafer framework. In particular, we refer to Shafer (1976a,b); Fagin and Halpern (1990); Gilboa and Schmeidler (1993); Coletti et al. (2016); Coletti and Vantaggi (2008).

Let Bel be a belief function on $(\Omega, \mathcal{F})$, Pl its dual plausibility function and $A \in \mathcal{F}, B \in \mathcal{F} \setminus \{\emptyset\}$.

General (Bayes) conditioning rule. The *general* conditional belief function of A given B, firstly recognized in Dempster (1967) and also called *Bayes* conditional belief function, is defined as

$$Bel^G(A|B) = \frac{Bel(A \cap B)}{Bel(A \cap B) + Pl(A^C \cap B)}, \tag{1.29}$$

where $Bel(A \cap B) + Pl(A^C \cap B) > 0$.

Its conjugate function, the *general (Bayes)* conditional plausibility function, has the same structure:

$$\begin{aligned} Pl^G(A|B) = 1 - Bel(A^C|B) = 1 - &\frac{Bel(A^C \cap B)}{Bel(A^C \cap B) + Pl(A \cap B)} \\ = &\frac{Pl(A \cap B)}{Pl(A \cap B) + Bel(A^C \cap B)}, \end{aligned} \tag{1.30}$$

where $Pl(A \cap B) + Bel(A^C \cap B) > 0$.

A *general (Bayes)* conditional belief (plausibility) function continues to be a belief (plausibility) function and it is the *lower (upper) envelope* of the set of all conditional probabilities defined by its core

$$Bel^G(A|B) = \min_{P\in \mathrm{core}(Bel)} P(A|B), \tag{1.31}$$

$$Pl^G(A|B) = \max_{P\in \mathrm{core}(Bel)} P(A|B), \tag{1.32}$$

where $P(\cdot|\cdot)$ is a conditional probability (see, also, Coletti et al., 2016). This property is also called *coherence* in the sense of Walley (Walley, 1982).

The *general* conditioning rule continues to hold in terms of a capacity ν and its dual function $\overline{\nu}$; if ν is a 2-monotone (alternating) capacity, the general conditional capacity continues to be 2-monotone (alternating) and equations (1.31)–(1.32) continue to hold (Denneberg, 1994).

If $A_1, A_2 \in \mathcal{F}$ and the conditional beliefs $Bel^G(A_i|B)$ for $i = 1, 2$ are defined, then $A_1 \subseteq A_2 \Rightarrow Bel^G(A_1|B) \leq Bel^G(A_2|B)$ (the property continues to hold if Bel reduces to a capacity).

Dempster conditioning rule. Proposed in terms of plausibility function Pl by Dempster (1967), the *Dempster conditioning rule* is defined, with $Pl(B) > 0$, as

$$Pl^D(A|B) = \frac{Pl(A \cap B)}{Pl(B)}. \tag{1.33}$$

Its conjugate belief function is

$$\begin{aligned} Bel^D(A|B) = 1 - Pl(A^C|B) = 1 - \frac{Pl(A^C \cap B)}{Pl(B)} \\ = \frac{Bel(A \cup B^C) - Bel(B^C)}{1 - Bel(B^C)}. \end{aligned} \tag{1.34}$$

The Dempster conditioning rule continues to hold in terms of a capacity ν and the Dempster conditional belief (plausibility) function Bel^D (Pl^D) continues to be a belief (plausibility) function.

The conditionals Bel^D and Pl^D are, respectively, the *lower* and the *upper envelope* of a set of conditional probabilities denoted as $\mathrm{core}(Bel^D(\cdot|B))$: starting from the conditional belief $Bel^D(\cdot|B)$ and taking its core

$$\mathrm{core}(Bel^D(\cdot|B)) = \{P \in \mathbf{P}(\Omega, \mathcal{F}) \ : \ P(A) \geq Bel^P(A|B), \ \forall A \in \mathcal{F}\}, \tag{1.35}$$

then $Bel^D(\cdot|B)$ and $Pl^D(\cdot|B)$ are, respectively, the lower and the upper envelope (Denneberg, 1994):

$$Bel^D(A|B) = \min_{P\in \mathrm{core}(Bel^D(\cdot|B))} P(A|B); \tag{1.36}$$

$$Pl^D(A|B) = \max_{P\in \mathrm{core}(Bel^D(\cdot|B))} P(A|B). \tag{1.37}$$

Note 2: *Conditions* (1.31)–(1.32) *say that the general conditional belief is the lower (upper) envelope of a set of probability measures, then all information have a relevant role. On the contrary, conditions* (1.36)–(1.37) *show that in the Dempster conditioning rule just the local information restricted to the conditioning event B have a role.*

Product conditioning rule. The product conditioning rule is analogous to the Dempster conditioning rule but it is defined in terms of a belief function, and it is also called *geometric* conditioning rule (Suppes and Zanotti, 1977; Shafer, 1976b) or *strong* conditioning rule (Planchet, 1989). The product conditioning rule is defined, with $Bel(B) > 0$, as

$$Bel^P(A|B) = \frac{Bel(A \cap B)}{Bel(B)}. \tag{1.38}$$

Its conjugate function $Pl^P(A|B)$ has the same structure of (1.34), hence

$$Pl^P(A|B) = 1 - Bel^P(A^C|B) = \frac{Pl(A \cup B^C) - Pl(B^C)}{1 - Pl(B^C)}. \tag{1.39}$$

A product conditional belief (plausibility) function Bel^P (Pl^P) continues to be a belief (plausibility) function.

The conditionals Bel^P and Pl^P are, respectively, the *lower* and the *upper envelope* of the core($Bel^P(\cdot|B)$) (see, e.g., Denneberg, 1994)

$$Bel^P(A|B) = \min_{P \in \text{core}(Bel^P(\cdot|B))} P(A|B); \tag{1.40}$$

$$Pl^P(A|B) = \max_{P \in \text{core}(Bel^P(\cdot|B))} P(A|B); \tag{1.41}$$

and an equivalent interpretation as that in Note 2 holds.

The following relations between rules hold:

$$Bel^G(A|B) \leq Bel^D(A|B) \leq Pl^D(A|B) \leq Pl^G(A|B); \tag{1.42}$$

$$Bel^G(A|B) \leq Bel^P(A|B) \leq Pl^P(A|B) \leq Pl^G(A|B). \tag{1.43}$$

Weak conditioning rule. Proposed by Planchet (1989) in terms of a belief function Bel on $(\Omega, \mathcal{F})$, the *weak* conditioning rule of A, given B with $Pl(B) \neq 0$, is defined by

$$Bel^W(A|B) = \frac{Bel(A) - Bel(A \cap B^C)}{1 - Bel(B^C)}. \tag{1.44}$$

The weak conditional Bel^W is still a belief function on $(\Omega, \mathcal{F})$ and its conjugate plausibility function is

$$Pl^W(A|B) = \frac{Pl(A) + Pl(B) - Pl(A \cup B)}{Pl(B)}. \tag{1.45}$$

The link between the *weak* conditioning rule and the *Dempster* conditioning rule is given by the following inequalities

$$Pl^W(A|B) \geq Pl^D(A|B), \quad Bel^W(A|B) \leq Bel^D(A|B). \tag{1.46}$$

In order to define a criterion for the choice of a specific conditioning rule, Yu and Arasta (1994) list a set of desirable properties and study which of these are satisfied by the proposed conditioning rules. For a conditional belief $Bel(\cdot|\cdot)$, the list of desirable properties is:

(1) $Bel(\cdot|B)$ remains a belief function of $(\Omega, \mathcal{F})$, for any $B \in \mathcal{F}$ with $Bel(B) \neq 0$;

(2) if Bel reduces to a probability measure P, then $Bel(\cdot|\cdot)$ reduces to a conditional probability;

(3) given $B \in \mathcal{F}$, with $Bel(B) \neq 0$, the set $A \in \mathcal{F}$ and $A \cap B$ have the same meaning, hence $Bel(A|B) = Bel(A \cap B|B)$;

(4) Bel is normalized: $Bel(B|B) = 1$ for any $B \in \mathcal{F}$ with $Bel(B) \neq 0$;

(5) $Bel(\cdot|\cdot)$ is commutative: $Bel(A|C|B) = Bel(A|B|C)$ for any $A, B, C \in \mathcal{F}$ and $Bel(B) \neq 0, Bel(C) \neq 0$;

(6) the conditioning rule allows repeated updating: $Bel(A|C|B) = Bel(A|B \cap C)$ for any $A, B, C \in \mathcal{F}$ with $Bel(B \cap C) \neq 0$;

(7) the conditional function satisfies the sandwich principle:
$\min(Bel(A|B), Bel(A|B^C)) \leq Bel(A) \leq \max(Bel(A|B), Bel(A|B^C))$ for $A, B \in \mathcal{F}$ with $Bel(B) \neq 0$ and $Bel(B^C) \neq 0$.

They prove that the *product* conditioning rule and the *Dempster* conditioning rule satisfy all properties except *(7)*; the *weak* conditioning rule satisfies properties *(1),(2),(5)*; the *general (Bayes)* conditioning rule satisfies properties *(1)-(4)* while *(5)-(6)* are satisfied just for some specific belief space.

The Dempster conditioning rule for belief functions records the effect of additional information, in fact, in terms of Möbius inverse, the masses of the focal events that meet both B and B^C are transferred into intersections between the event and the conditioning one; formally, denoting by $\mathcal{B} = \{A \in \mathcal{F} : \mu(A) \neq 0\}$ the set of focal events of Bel, then $\mu(A)$, with $A \in \mathcal{B}$, such that $A \cap B \neq \emptyset$ and $A \cap B^C \neq \emptyset$, is transferred to events $A \cap B$.

On the contrary, the general (Bayes) conditioning rule selects the relevant information and, in terms of Möbius inverse, the masses of focal events that meet both B and B^C are transferred to subset larger that $A \cap B$, in particular, given $A \in \mathcal{B}$ such that $A \cap B \neq \emptyset$ and $A \cap B^C \neq \emptyset$, then $\mu(A)$ is transferred to subsets of $C \cap B$ that are only larger than $A \cap B$, where $C = \cup_{A \in \mathcal{B}}$. Jaffray (2008) gives a way to directly compute $\mu^G(\cdot|B)$ from μ (for details, see Theorem 1 in Jaffray, 2008).

In the following example, inspired by Jaffray (2008), we show how the general (Bayes), Dempster and product conditioning rules work.

Example 1.3 *Consider $\Omega = \{1, 2, 3\}$, $\mathcal{F} = \mathcal{P}(\Omega)$ and the lower probability given in Example 1.1, that is a belief function since $\mu(A) \geq 0$ for all $A \in \mathcal{F}$. We report it, denoted as Bel, and its Möbius inverse (to avoid cumbersome notation we denote events omitting commas).*

$\mathcal{F}$	$\emptyset$	$\{1\}$	$\{2\}$	$\{3\}$	$\{12\}$	$\{13\}$	$\{23\}$	Ω
Bel	0	$\frac{1}{3}$	0	0	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{2}{3}$	1
μ	0	$\frac{1}{3}$	0	0	0	0	$\frac{2}{3}$	0

We compute the conditional belief functions with respect to the event $B = \{12\}$ *with* ***(i)*** *the general (Bayes) conditioning rule* $Bel^G(\cdot|B)$ *(and its Möbius inverse* μ^G*), with* ***(ii)*** *the Dempster conditioning rule* $Bel^D(\cdot|B)$ *(and its Möbius inverse* μ^D*) and with* ***(iii)*** *the product conditioning rule* $Bel^P(\cdot|B)$ *(and its Möbius inverse* μ^P*).*

$\mathcal{F} \setminus \{\emptyset\}$	$\{1\}$	$\{2\}$	$\{3\}$	$\{12\}$	$\{13\}$	$\{23\}$	Ω
Bel	$\frac{1}{3}$	0	0	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{2}{3}$	1
μ	$\frac{1}{3}$	0	0	0	0	$\frac{2}{3}$	0
$Bel^G(\cdot\|B)$	$\frac{1}{3}$	0	0	1	$\frac{1}{3}$	0	1
μ^G	$\frac{1}{3}$	0	0	$\frac{2}{3}$	0	0	0
$Bel^D(\cdot\|B)$	$\frac{1}{3}$	$\frac{2}{3}$	0	1	$\frac{1}{3}$	$\frac{2}{3}$	1
μ^D	$\frac{1}{3}$	$\frac{2}{3}$	0	0	0	0	0
$Bel^P(\cdot\|B)$	1	0	0	1	1	0	1
μ^P	1	0	0	0	0	0	0

Note that $\mu^i \geq 0$ *for* $i = \{\text{“}G\text{”}, \text{“}D\text{”}, \text{“}P\text{”}\}$*, then the conditional functions continue to be belief functions. In particular,* $Bel^D(\cdot|B)$ *and* $Bel^P(\cdot|B)$ *reduce to a probability measure. We stress that it is a particular case due to the values of the example. In this case, since their* core *reduces to a singleton*

$$\begin{aligned} \mathrm{core}(Bel^D(\cdot|B)) &= \left\{Bel^D(\cdot|B)\right\} = \left\{P^D\right\}, \\ \mathrm{core}(Bel^P(\cdot|B)) &= \left\{Bel^P(\cdot|B)\right\} = \left\{P^P\right\}, \end{aligned}$$

(1.36) *and* (1.40) *are automatically satisfied.*

On the contrary, we investigate if the general conditioning rule is coherent, i.e., it is the lower envelope of the set $\{P(A|B) \ : \ P \in \mathrm{core}(Bel), \ \forall A \in \mathcal{F}\}$.

The set of extreme points $\mathrm{ext}(\mathrm{core}(Bel)) = \{P^\sigma \ : \ \sigma \in \Sigma\}$*, where* $\Sigma = \{1, 2, 3\}$*, is the following*

$$\begin{aligned} P^{(1,2,3)} &= \left(\tfrac{1}{3}, 0, \tfrac{2}{3}\right) = P_1, & P^{(1,3,2)} &= \left(\tfrac{1}{3}, \tfrac{2}{3}, 0\right) = P_2, \\ P^{(2,1,3)} &= \left(\tfrac{1}{3}, 0, \tfrac{2}{3}\right) = P_1, & P^{(2,3,1)} &= \left(\tfrac{1}{3}, 0, \tfrac{2}{3}\right) = P_1, \\ P^{(3,1,2)} &= \left(\tfrac{1}{3}, \tfrac{2}{3}, 0\right) = P_2, & P^{(3,2,1)} &= \left(\tfrac{1}{3}, \tfrac{2}{3}, 0\right) = P_2, \end{aligned}$$

with respect to which we compute the conditional probabilities and the lower conditional probability $\underline{P}(\cdot|B) = \min_{P \in \mathrm{ext}(\mathrm{core}(Bel))} P(\cdot|B)$.

$\mathcal{F}\setminus\{\emptyset\}$	$\{1\}$	$\{2\}$	$\{3\}$	$\{12\}$	$\{13\}$	$\{23\}$	Ω
P_1	$\frac{1}{3}$	0	$\frac{2}{3}$	$\frac{1}{3}$	1	$\frac{2}{3}$	1
$P_1(\cdot\|B)$	1	0	0	1	1	0	1
P_2	$\frac{1}{3}$	$\frac{2}{3}$	0	1	$\frac{1}{3}$	$\frac{2}{3}$	1
$P_2(\cdot\|B)$	$\frac{1}{3}$	$\frac{2}{3}$	0	1	$\frac{1}{3}$	$\frac{2}{3}$	1
$\underline{P}(\cdot\|B)$	$\frac{1}{3}$	0	0	1	$\frac{1}{3}$	0	1

Since $Bel^G(A|B) = \min_{P\in\mathrm{core}(Bel)} P(A|B)$, *for all* $A \in \mathcal{F}$, *then* $Bel^G(\cdot|B)$ *is coherent. Also, relations between conditioning rule given in* (1.42)–(1.43) *hold.*

Note that, as explained before, the Dempster conditioning rule transfers the masses of events that meet both B *and* B^C *to events* $A \cap B$. *In this the events that meet both* B *and* B^C *is only* $A = \{23\}$ *with mass* $\mu(\{23\}) = \frac{2}{3}$. *Since we have* $A \cap B = \{23\} \cap \{12\} = \{2\}$, *the mass* $\mu(\{23\})$ *is transferred to event* $\{2\}$, *in fact* $\mu^{DS}(\{2\}) = \frac{2}{3}$.

On the contrary, the general (Bayes) conditioning rule transfers the mass of A *to events larger than* $A \cap B$ *that are subsets of* $C \cap B$. *In this example, events larger than* $\{2\}$ *that are subsets of* $C \cap B = \Omega \cap \{12\} = \{12\}$ *is only* $\{12\}$ *itself. In fact* $\mu^G(\{12\}) = \frac{2}{3}$.

♦

Arguments for belief functions

The theory of belief function firstly proposed by Dempster and Shafer provides the foundation for the majority of this thesis. The selection of belief functions over other non-additive measures, in particular 2-monotone capacities that share the lower envelope property (1.21) and the lower expectation property (1.27), is due to the clear interpretability and the lower computational cost. The Möbius inverse of a belief function can be easily interpreted as a measure of evidence in favour of a set, as Examples 1.1–1.2 show. On the other side, the 2-monotone framework does not require μ to be non-negative, hence, from a "real-world" point of view, it may be more difficult to interpret. For a deeper discussion see, e.g., Cuzzolin (2021); Shafer (1990).

Multivalued mapping. For finite Θ and Ω, given a probability space $(\Theta, 2^\Theta, P)$ and a *multivalued mapping* $\Gamma : \Theta \to 2^\Omega$, they induce a belief function $Bel : 2^\Omega \to [0,1]$ computing for each $\theta \in \Theta$ the degree of belief of an event $A \subseteq \Omega$ as the total probability of all θ

$$Bel(A) = \sum_{\theta\in\Theta|\Gamma(\theta)\subseteq A} P(\theta). \tag{1.47}$$

In other words, the Möbius inverse $\mu : 2^\Omega \to [0,1]$ is given by $\mu(\cdot) = P(\Gamma = \cdot)$. Conditioning to an event $H \in 2^\Omega$, we have to consider $H_* = \{\theta \in \Theta \ : \ \Gamma(\theta) \subseteq H\}$ and the following relation holds

$$P(\cdot|H_*) = \frac{P(\cdot \cap H_*)}{P(H_*)} = \frac{\mu(\cdot \cap H)}{Bel(H)}. \tag{1.48}$$

Generalized non-additive probabilities. As characterized in Section 1.2.1, belief functions are particular non-additive measures such that they are the lower envelope of their core (see (1.21)) and the Choquet integral with respect to Bel is the lower expectation among each $P \in \text{core}(Bel)$ (see (1.27)). 2-monotone capacities also satisfies these properties.

Inner measures. Given a measurable space $(\Omega, \mathcal{F})$, let us suppose to have a probability measure P defined over a σ-field of subsets $\mathcal{I}$ of $\mathcal{F}$. The inner probability of P is such an extension of P to $\mathcal{F}$ and it measures the degree we should believe in an event for which the probability P is not defined. The inner measure P_* is defined, for all $A \in \mathcal{F}$, as

$$P_*(A) = \max\{P(B) \ : \ B \subseteq A, \, B \in \mathcal{I}\}, \tag{1.49}$$

and it is proved to be a belief function (the converse, in general, is not true, see, e.g., Fagin and Halpern, 1991).

Chapter 2

Arbitrage theory in discrete time models

In this Chapter, we look at the fundamental theorems and models that are based on classical no-arbitrage assumptions in a one-period setting (Section 2.1) and then we consider the generalization to a multi-period setting (Section 2.2). The absence of arbitrage opportunities, that is the absence of a chance to gain without risk, is a reasonable and necessary assumption in the primary literature that leads to fundamental theorems on which pricing is based. Finally, the binomial and trinomial market models are investigated (Sections 2.2.1–2.2.2).

2.1 One-period setting

Consider a financial market open at times $t = 0$ and $t = 1$. The market is composed by a set of *assets* (or securities), each defined as a contract between two investors which specifies, for each state of the world $i \in \Omega$, an amount (positive or negative) of money or commodity that the seller of the contract has to transfer to the buyer.

The cash flow is deterministic (i.e., it does not depend on future states of the world) when the asset is *riskless*, otherwise the cash flow is a random variable depending on which state of the world will occur and it is called *risky*. An example of riskless asset is the risk-free bond: an asset that gives to the holder the right to receive a predetermined amount of money in each future state of the world. The rate of return of a risk-free bond is called *risk-free interest rate*.

The financial market is commonly assumed to satisfy the following fundamental assumptions (Allingham, 1991; Dybvig and Ross, 1989):

(i) absence of frictions (there are no transaction costs, taxes and other restrictions on trading);

(ii) competitiveness (every quantity can be traded at market's price).

A one-period market model is composed by a set of K risky assets, denoted by $S^1, \ldots, S^K$, and by a riskless asset (bond), denoted by S^0. Risky assets have a price process denoted as $\left\{S_0^k, S_1^k\right\}$, for $k = 1, \ldots, K$, where $S_0^k = s_0^k > 0$ is a deterministic value called *price*, and S_1^k is a random variable depending on which state of the world will occur, called *payoff*. The riskless bond S^0, without loss of generality, has price process

$$\left\{S_0^0 = 1,\ S_1^0 = (1+r)S_0^0 = (1+r)\right\}, \tag{2.1}$$

where $r > -1$ is the risk-free interest rate of the market. Its price is usually used as a numéraire (see, e.g., Pliska, 1997), meaning that it allows to discount risky price processes and to define a new process called *discounted price process* which we denote by $\left\{\tilde{S}_0^k, \tilde{S}_1^k\right\}$, where $\tilde{S}_0^k = S_0^k$ and $\tilde{S}_1^k = \frac{S_1^k}{S_1^0} = (1+r)^{-1}S_1^k$, for $k = 1, \dots, K$.

The set of assets of the market is denoted as a process $\{\boldsymbol{S}_0, \boldsymbol{S}_1\}$ on $\mathbb{R}^{K+1}$ where $\boldsymbol{S}_0 = \left(S_0^0, \dots, S_0^K\right)$ is the vector of prices and $\boldsymbol{S}_1 = \left(S_1^0, \dots, S_1^K\right)$ is the vector of payoffs.

Price processes are defined on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_0, \mathcal{F}_1\}, P)$ where $\Omega = \{1, \dots, n\}$, with $n \in \mathbb{N}$, is a finite state space, $\{\mathcal{F}_0, \mathcal{F}_1\}$ is a filtration such that $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and $\mathcal{F}_1 = \mathcal{P}(\Omega) = \mathcal{F}$ is the power set of Ω, and P is a probability measure on $\mathcal{F}$, with $P(\{i\}) > 0$, that can be interpreted as the *"real-world"* probability (also called *natural* probability).

Note 3: *In classical finite-state no-arbitrage pricing models (see, e.g., Delbaen and Schachermayer, 2006; Pliska, 1997), the positivity of the "real-world" probability P is motivated by the fact that only "realistic" states of nature are taken into account, i.e., states with null measure are discarded; this assures that an asset with non-negative and non-null payoff will have a positive price at time $t = 0$. Furthermore, a common assumption in finance is that all market agents share the same probabilistic opinions, i.e., they have the same P.*

In agreement with Chapter 1, we denote by $\mathbb{R}^\Omega$ the set of all $\mathcal{F}$-measurable random variables and by $\mathbf{P}(\Omega, \mathcal{F})$ the set of all probability measures on $(\Omega, \mathcal{F})$.

Definition 2.1 (Price assessment) *A function $\pi' : \boldsymbol{S}_1 \to \mathbb{R}$ is called price assessment if and only if $\pi'(S_1^k) = S_0^k$, for all $k = 0, \dots, K$.*

We suppose that the price assessment function $\pi'(\cdot)$ is given.

Definition 2.2 (One-period portfolio) *A portfolio (or trading strategy) is a collection of assets that an agent can hold. It is denoted by a vector $\boldsymbol{\lambda} = (\lambda^0, \dots, \lambda^K) \in \mathbb{R}^{K+1}$ whose k-th component λ^k expresses the number of units purchased ($\lambda^k > 0$) or sold ($\lambda^k < 0$) of the k-th asset in the one-period time interval $[0, 1]$.*

At time $t = 0$ the price of the portfolio $\boldsymbol{\lambda} \in \mathbb{R}^{K+1}$ is computed as the weighted sum of prices and it is denoted by

$$V_0 = \boldsymbol{\lambda}^T \boldsymbol{S}_0 = \sum_{k=0}^{K} \lambda^k S_0^k = \sum_{k=0}^{K} \lambda^k \pi'(S_1^k). \tag{2.2}$$

At time $t = 1$ the payoff of the portfolio $\boldsymbol{\lambda} \in \mathbb{R}^{K+1}$ is a random variable $V_1 : \Omega \to \mathbb{R}$ defined, for every $i \in \Omega$, as

$$V_1 = \boldsymbol{\lambda}^T \boldsymbol{S}_1 = \sum_{k=0}^{K} \lambda^k S_1^k. \tag{2.3}$$

2.1.1 Dutch-book and arbitrage opportunities

"An arbitrage opportunity is an investment strategy that guarantees a positive payoff in some contingency with no possibility of a negative payoff and with no net investment. By assumption, it is possible to run the arbitrage possibility at arbitrary scale; in other words, an arbitrage opportunity represents a money pump. A simple example of arbitrage is the opportunity to borrow and lend costlessly at two different fixed rates of interest. Such a disparity between the two rates cannot persist: arbitrageurs will drive the rate together." (Dybvig and Ross, 1989, p.57)

Definition 2.3 (Dutch-book) *A portfolio* $\boldsymbol{\lambda} \in \mathbb{R}^{K+1}$ *avoids a Dutch-book opportunity if the following condition holds*

$$\max_{i \in \Omega} \sum_{k=0}^{K} \lambda^k \left(\tilde{S}_1^k(i) - \pi'(S_1^k) \right) \geq 0. \tag{2.4}$$

The no-Dutch-book portfolio is also called *coherent* since it does not allow a sure loss for the holder. On the contrary, a portfolio $\boldsymbol{\lambda} \in \mathbb{R}^{K+1}$ that allows a Dutch-book opportunity is called *incoherent* (Walley, 1982, 1991).

The concept of arbitrage opportunity is stronger than the Dutch-book one, since an arbitrage portfolio allows to acquire a positive amount of money without risk, also called *free lunch*. The existence of an arbitrage portfolio allows agents to acquire an unlimited amounts of money without risk, called *money pump*, so it is a possibility that has to be forbidden.

Definition 2.4 (One-period arbitrage opportunity) *A portfolio* $\boldsymbol{\lambda} \in \mathbb{R}^{K+1}$ *is an arbitrage portfolio if it guarantees a positive payoff with a non-positive price, or the opposite, i.e., it satisfies one of the following two conditions:*

(i) $V_0 < 0$ *and* $V_1 = 0$*;*

(ii) $V_0 \leq 0$ *and* $V_1 \geq 0$ *with* $V_1 \neq 0$ *(i.e., there is a strict inequality for at least one* $i \in \Omega$*), equivalently written as*

$$\sum_{k=0}^{K} \lambda^k \left(\tilde{S}_1^k(i) - \pi'(S_1^k) \right) \geq 0, \tag{2.5}$$

for all i*, with a strict inequality for at least one* $i \in \Omega$*.*

A Dutch-book opportunity implies the existence of an arbitrage opportunity but the converse does not hold as showed in the following example from Schervish et al. (2008).

Example 2.1 *Let be* $\Omega = \{0, 1\}$*,* X *a random variable that takes values* $X(0) = 0$*,* $X(1) = 1$*. Let be* $\pi'(X) = 0$*. For* $\lambda = 1$*, we have that*

$$\max\{\lambda(X(0) - \pi'(X)), \lambda(X(1) - \pi'(X))\} = \max\{0, 1\} = 1,$$

hence, condition (2.4) *is satisfied, that is this portfolio avoids a Dutch-book opportunity. Nevertheless, condition* (2.5) *is satisfied, then it is an arbitrage opportunity.*

♦

The assumption that arbitrage opportunities must not exist stems from individual rationality since a rational agent who prefers more to less will maintain an arbitrage position in the absence of a scale constraint. For this reason, pricing theory commonly makes the assumption that the market must be *arbitrage-free* (see, e.g., Pliska, 1997; Pascucci and Runggaldier, 2011; Dybvig and Ross, 1989; Bjork, 2009).

The following fundamental theorems demonstrate how the assumption of no-arbitrage has significant implications.

2.1.2 Fundamental theorems

Theorem 2.1 (Fundamental theorem of asset pricing (Dybvig and Ross, 1989)) The following statements are equivalent:

(i) absence of arbitrage opportunities;

(ii) existence of a consistent positive linear pricing rule;

(iii) existence of an optimal demand for some agent who prefers more to less.

Condition *(iii)* is not pointed out in this work, then we refer to Dybvig and Ross (1989) for details. In what follows, we continue to assume a finite setting, in agreement with Chapter 1.

Definition 2.5 (Pricing rule) *A pricing rule (or price functional) is a function, denoted by $\pi : \mathbb{R}^\Omega \to \mathbb{R}$, that assigns a value to all possible random payoffs $X \in \mathbb{R}^\Omega$. A pricing rule is linear if, given two random variables $X, Y \in \mathbb{R}^\Omega$, it holds that $\pi(\alpha X + \beta Y) = \alpha\pi(X) + \beta\pi(Y)$, with $\alpha, \beta \in \mathbb{R}$.*

The pricing rule has to be *consistent* with the marketed assets, i.e., it has to be such that, for all $k = 0, \ldots, K$, the following equality holds

$$S_0^k = \pi(S_1^k) = \pi'(S_1^k). \tag{2.6}$$

Note 4: *The pricing rule $\pi(\cdot)$ is an extension of the price assessment $\pi'(\cdot)$ defined on $\boldsymbol{S}_1$, to all assets, also hypothetical, defined over the same set of states.*

A positive linear pricing rule can be represented in alternative ways that could be more or less useful depending on the context. This will be discussed after the next fundamental theorem.

Theorem 2.2 (Pricing rule representation theorem (Dybvig and Ross, 1989)) The following statements are equivalent:

(a) existence of a positive linear pricing rule;

(b) existence of positive risk-neutral probabilities and an associated riskless rate (*the martingale property*);

(c) existence of a positive state price density.

Approaches *(b)* and *(c)* are equivalent but they can be more or less preferred depending on the framework of study. In particular, the risk-neutral approach *(b)* is useful for valuation's problems without going through individual preferences since we ignore individual utilities and beliefs expressed by P; the state price density *(c)* is preferred if we want to study choice problems. We will not focus on the state price density, for details we refer to Dybvig and Ross (1989).

The Pricing Rule Representation Theorem assures that the positive linear pricing rule π is equivalent to the existence of a probability $Q \in \mathbf{P}(\Omega, \mathcal{F})$ called *risk-neutral*, such that:

(1) it is equivalent to the natural probability (in symbol, $Q \sim P$): for every $A \in \mathcal{F}$, it holds that $P(A) = 0 \Leftrightarrow Q(A) = 0$;

(2) the price assessment is a discounted expected value:

$$\pi'(S_1^k) = (1+r)^{-1}\mathbb{E}_Q(S_1^k) = S_0^k, \tag{2.7}$$

for $k = 1, \dots, K$. It can be equivalently written in terms of discounted price process in the following way: $\pi'(S_1^k) = \mathbb{E}_Q\left(\tilde{S}_1^k\right) = S_0^k$. In this way we can note that the discounted process is a *martingale* with respect to Q; for this reason Q is also called *equivalent martingale measure.*

The no-arbitrage assumption assures that there exists a pricing rule given by the discounted expected value of payoffs which extends the price assessment π', defined over the marketable assets $\boldsymbol{S}_1$, to a pricing rule π over the whole set $\mathbb{R}^\Omega$.

This result is stated in the first fundamental theorem of asset pricing that summarizes Theorem 2.1 and Theorem 2.2. By a financial point of view, the next theorem has the following interpretation: the no-arbitrage assumption asks that there could not be sure gains generated by buying and short-selling assets, then the price process cannot be increasing on average but it has to be constant on average (i.e., a martingale process).

Theorem 2.3 (First fundamental theorem of asset pricing)
The following statements are equivalent:

(i) absence of arbitrage opportunities;

(ii) existence of an equivalent martingale measure.

The first fundamental theorem of asset pricing does not assure the *uniqueness* of the equivalent martingale measure. In general, it assures that there exists a *class* of equivalent martingale measures defined as

$$\mathcal{Q} = \{Q \in \mathbf{P}(\Omega, \mathcal{F}) : S_0^k = (1+r)^{-1}\mathbb{E}_Q(S_1^k),\ k = 1, \dots, K\}. \tag{2.8}$$

Suppose that X is a *European-type derivative*, that is a financial contract defined over the filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_0, \mathcal{F}_1\}, P)$ with price process $\{X_0, X_1\}$, where the payoff at maturity $T = 1$ is a random variable $X_1(i)$ depending on which $i \in \Omega$ occurs, and X_0 is the price of the derivative at time $t = 0$. The derivative is called *replicable* if there exists a portfolio $\boldsymbol{\lambda} \in \mathbb{R}^{K+1}$ (called *replicating portfolio*) such that the derivative and the replicating portfolio have the same payoff: $X_1 = V_1$.

If the market is free of arbitrage opportunities the *law of one price* holds: the value of a replicable derivative is the same of its replicating portfolio

$$X_0 = V_0 = \sum_{k=0}^{K} \lambda^k S_0^k. \tag{2.9}$$

Theorem 2.4 (Risk neutral valuation principle (Pliska, 1997))
If the one-period market model is arbitrage-free, then the price of a replicable derivative X is the discounted expected value of its payoff

$$X_0 = (1+r)^{-1}\mathbb{E}_Q(X_1), \tag{2.10}$$

where Q is any equivalent martingale measure in $\mathcal{Q}$. Hence, under the no-arbitrage hypothesis, a pricing rule is expressed as $\pi(\cdot) = (1+r)^{-1}\mathbb{E}_Q(\cdot)$.

Remark 1. Theorem 2.4 does not guarantee that the price X_0 is unique, since different $Q \in \mathcal{Q}$ could lead to different prices, each of them consistent with the no-arbitrage assumption.

Definition 2.6 (Completeness) *A market model in which each derivative X is replicable is called complete.*

We introduce the matrix notation to deal with *incompleteness*. Payoffs of the riskless asset and of K risky assets, for each $i \in \Omega$, are defined in the matrix $A \in \mathbb{R}^{n\times(K+1)}$

$$A = \begin{bmatrix} S_1^0(1) & S_1^1(1) & \cdots & S_1^K(1) \\ \vdots & \vdots & & \vdots \\ S_1^0(n) & S_1^1(n) & \cdots & S_1^K(n) \end{bmatrix}, \tag{2.11}$$

and the vector of payoffs of the derivative is denoted by $\boldsymbol{X} = (X_1(1), \dots, X_1(n)) \in \mathbb{R}^n$.

The next theorem links the completeness of the market (that is equivalent to say that there exists a replicating portfolio for each derivative) with the uniqueness of the equivalent martingale measure.

Theorem 2.5 (Second fundamental theorem of asset pricing)
If the market model is free of arbitrage opportunities, the following statements are equivalent:

(i) there exists a portfolio $\boldsymbol{\lambda} \in \mathbb{R}^{K+1}$ such that the linear problem $A\boldsymbol{\lambda}^T = \boldsymbol{X}$ has a unique solution, for all $X \in \mathbb{R}^{\Omega}$. It occurs if and only if $rank(A) = n = K+1$ (supposing that there are no redundant assets);

(ii) the market is complete;

(iii) the set of equivalent martingale measures $\mathcal{Q}$ reduces to a singleton $\mathcal{Q} = \{Q\}$.

Note 5: *The importance of the completeness of a market comes from the fact that it allows to price and hedge all possible contingent claims, also hypothetical, in a preference-independent way.*

On the contrary, if the market is incomplete, each equivalent martingale measure in the set $\mathcal{Q}$ leads to a different prices for X, as said in Remark 1. Anyhow, the fair price of X_1 (i.e., the price that does not give rise to an arbitrage opportunity) is proved to be in the interval $\Big(\underline{V}(X_1), \overline{V}(X_1)\Big)$, defined through the closest replicable derivative of X (Pliska, 1997)[1].

The simplest approach to overcome market incompleteness is to *complete* the market; other techniques look for the best replicating strategy and they will be discussed in Chapter 3.

2.2 Multi-period setting

A multi-period financial market is a market with $T \in \mathbb{N}$ finite trading dates $t = 0, 1, \ldots, T$, defined on a finite filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t=0}^T, P)$.

As the one-period market model, we consider a multi-period market composed by K risky assets with non negative price process $\{S_t^k\}_{t=0}^T$, for $k = 1, \ldots, K$, and by one riskless asset with price process $\{S_t^0\}_{t=0}^T$ such that

$$S_0^0 = 1, \quad S_t^0 = (1+r)^t, \tag{2.12}$$

where $r > -1$ is the *constant* risk-free interest rate of the market in the time interval $[t, t+1]$.

Price processes are adapted to the filtration $\{\mathcal{F}_t\}_{t=0}^T$ with $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and $\mathcal{F}_T = \mathcal{P}(\Omega) = \mathcal{F}$, such that $\mathcal{F}_t \subset \mathcal{F}_{t+1}$ for every $t = 0, \ldots, T-1$; each algebra $\mathcal{F}_t$ is generated by a unique partition $\mathcal{P}_t$ of Ω and, at time t, agents know which $A \in \mathcal{P}_t$ contains the true state of the world[2].

Definition 2.7 (Multi-period portfolio) *In a multi-period setting with times $t = 0, \ldots, T$, a portfolio (or trading strategy) is a random process $\{\boldsymbol{\lambda}_t\}_{t=0}^{T-1}$ such that, at time t, it is a random vector $\boldsymbol{\lambda}_t = (\lambda_t^0, \lambda_t^1, \ldots, \lambda_t^K) \in \mathbb{R}^{K+1}$, where $\lambda_t^0 > 0$ ($\lambda_t^0 < 0$) is the bond's units bought (sold) in the time interval $[t, t+1]$, and $\lambda_t^k > 0$ ($\lambda_t^k < 0$) is the number of units of the k-th asset bought (sold) in the time interval $[t, t+1]$.*

We stress that the trading strategy is set up to time $T-1$ and that the trading strategy $\boldsymbol{\lambda}_t$ has to be measurable with respect to $\mathcal{F}_t$ and constant on the partition $\mathcal{P}_t$.

The value of the portfolio is a random process $\{V_0, \ldots, V_T\}$ where V_t is the value on the t-th one-period interval $[t, t+1]$, for $t = 0, \ldots, T-1$, computed with respect to the strategy $\boldsymbol{\lambda}_t$ settled at time t, and it is denoted as

$$V_t = \sum_{k=0}^{K} \lambda_t^k S_t^k = \lambda_t^0 (1+r)^t + \sum_{k=1}^{K} \lambda_t^k S_t^k, \tag{2.13}$$

[1] If Z_1 is the payoff of the closest replicable derivative of X, bound are computed in the following way: $\overline{V}(X_1) = \inf_{\substack{Z_1 \leq X_1, \\ Z_1 \text{ is replicable}}} (1+r)^{-1} \mathbb{E}_Q(Z_1)$, $\underline{V}(X_1) = \sup_{\substack{Z_1 \leq X_1, \\ Z_1 \text{ is replicable}}} (1+r)^{-1} \mathbb{E}_Q(Z_1)$.

[2] This feature will be later called *completely resolving uncertainty.*

while, for $t = T$, the value of the portfolio in the one-period interval is given by

$$V_T = \sum_{k=0}^{K} \lambda_{T-1}^k S_T^k = \lambda_{T-1}^k (1+r)^T + \sum_{k=1}^{K} \lambda_{T-1}^k S_T^k. \tag{2.14}$$

Once that information at time t are revealed, the investor changes his strategy from $\boldsymbol{\lambda}_t$ to $\boldsymbol{\lambda}_{t+1}$ and the value of the portfolio immediately becomes

$$\lambda_{t+1}^0 (1+r)^{t+1} + \sum_{k=1}^{K} \lambda_{t+1}^k S_{t+1}^k. \tag{2.15}$$

Definition 2.8 (Self-financing portfolio) *A portfolio $\{\boldsymbol{\lambda}_t\}_{t=0}^{T-1}$ is self-financing if, for $t = 1, \dots, T-1$, it satisfies the following equality*

$$V_t = \lambda_{t-1}^0 (1+r)^t + \sum_{k=1}^{K} \lambda_{t-1}^k S_t^k = \lambda_t^0 (1+r)^t + \sum_{k=1}^{K} \lambda_t^k S_t^k. \tag{2.16}$$

In words, a self-financing portfolio is such that the investor does not add or withdraw founds from the value of the portfolio at any of the times. We stress that in the one-period setting each portfolio is vacuously self-financing.

Considering the discounted price process of the *k-th asset* in t

$$\tilde{S}_t^k = \frac{S_t^k}{S_t^0} = (1+r)^{-t} S_t^k, \tag{2.17}$$

then the discounted price process of the *portfolio* is given by

$$\tilde{V}_t = \lambda_t^0 + \sum_{k=1}^{K} \lambda_t^k \tilde{S}_t^k = V_t (1+r)^{-t}, \tag{2.18}$$

for $t = 0, \dots, T-1$. For $t = T$ it holds that

$$\tilde{V}_T = \lambda_{T-1}^0 + \sum_{k=1}^{K} \lambda_{T-1}^k \tilde{S}_T^k = V_T (1+r)^{-T}. \tag{2.19}$$

The no-arbitrage principle and the fundamental theorems of asset pricing defined in the one-period setting can be straightforwardly extended to the multi-period setting as follows.

Definition 2.9 (Multi-period arbitrage opportunity) *An arbitrage opportunity in multi-period setting is a self-financing strategy $\{\boldsymbol{\lambda}_t\}_{t=0}^{T-1}$ such that it satisfies one of the following two conditions, where comparisons are intended over Ω:*

(i) $V_0 < 0$ *and* $V_T = 0$;

(ii) $V_0 \leq 0$ *and* $V_T \geq 0$ *with* $V_T \neq 0$.

Theorem 2.6 (Multi-period first fundamental theorem of asset pricing (Pliska, 1997))
In a multi-period market model there are no arbitrage opportunities if and only if there exists an equivalent probability measure $Q \sim P$ such that the discounted price process $\{\tilde{S}^k\}_{t=0}^T$ is a martingale under Q, meaning that, for every k and $1 \leq t \leq T, 0 \leq s \leq T-t$,

$$\mathbb{E}_Q(\tilde{S}^k_{t+s}|\mathcal{F}_t) = \tilde{S}^k_t. \tag{2.20}$$

The link between the one-period setting and the multi-period one is given by the following theorem.

Theorem 2.7 (Equivalence one-period multi-period arbitrage (Pliska, 1997))
If the multi-period model does not allow any arbitrage opportunity, then none of the underlying one-period models has any arbitrage opportunities in the single period sense.

Also, the *law of one price*, the definition of completeness and the second fundamental theorem of asset pricing continue to hold in multi-period setting.

2.2.1 Binomial model

The simplest example of market model showing all peculiarities of the no-arbitrage theory just discussed is the binomial market model proposed by Cox et al. (1979).

The binomial model is a discrete time model in which the market is open at times $t = 0, 1, \ldots, T$, defined over a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t=0}^T, P)$, where $\Omega = \{1, \ldots, 2^T\}$, composed by a riskless bond S^0 and a risky asset S^1.

The bond has price process

$$S^0_0 = 1, \quad S^0_t = (1+r)^t, \tag{2.21}$$

for $t = 1, \ldots, T$, where $r > 0$ is the constant risk-free interest rate of the market.

The risky asset has price process

$$S^1_0 = s_0 > 0, \quad S^1_t = \begin{cases} uS^1_{t-1} \text{ with probability } p \\ dS^1_{t-1} \text{ with probability } (1-p) \end{cases} \tag{2.22}$$

for $t = 1, \ldots, T$, where $u > d > 0$ are, respectively, the "up" and the "down" parameters, and $p \in (0,1)$ is the probability of an up movement.

The price process of the risky asset is characterized by a *natural* probability measure $P \in \mathbf{P}(\Omega, \mathcal{F})$ such that the probability of S^1 follows a binomial probability distribution

$$P(S^1_t = u^j d^{t-j} s_0) = \binom{t}{j} p^j (1-p)^{t-j}, \tag{2.23}$$

for $j = 0, \ldots, t$. The price process $\{S^1_t\}_{t=0}^T$ is Markovian and time-homogeneous: this is due to independence of log-returns over the single periods.

Theorem 2.8 (Pascucci and Runggaldier, 2011)
The binomial market model is free of arbitrage opportunities if and only if

$$u > (1+r) > d. \tag{2.24}$$

It is equivalent to the existence of a unique equivalent martingale measure $Q \sim P$, that is given by the risk-neutral parameter

$$q = \frac{(1+r)-d}{u-d}. \tag{2.25}$$

Note 6: *A Markovian random process with respect to a probability measure P may not continue to be Markovian with respect to an equivalent probability measure $Q \sim P$. However, in this setting, Pliska (1997) proves that if the market model is free of arbitrage opportunities, the discounted price process $\{\tilde{S}_t\}_{t=0}^T$ is Markovian with respect to P and the filtration $\mathcal{F}$ is the one generated by $\{\tilde{S}_t\}_{t=0}^T$, then there exists a martingale measure Q under which $\{\tilde{S}_t\}_{t=0}^T$ is a Markov chain.*

It follows that the discounted price process is a martingale with respect to Q, that is, for every $t = 1, \dots, T$, it holds that

$$\mathbb{E}_Q\left(\tilde{S}_t^1 | \mathcal{F}_{t-1}\right) = \tilde{S}_{t-1}^1. \tag{2.26}$$

Moreover, the equivalent martingale measure Q determines a binomial probability distribution for the risky asset

$$Q\left(S_t^1 = u^j d^{t-j} s_0\right) = \binom{t}{j} q^j (1-q)^{t-j}, \tag{2.27}$$

for $j = 0, \dots, t$.

Hence, condition (2.24) assures that the binomial market model is arbitrage-free and complete. Given a European-type derivative X, whose value depends on the underlying asset S^1, its payoff is $X_T = \varphi(S_T^1)$, where $\varphi(\cdot) : \mathbb{R} \to \mathbb{R}$ is a suitable *contract function*, and its value (price) at time t, for $t = 0, \dots, T-1$, can be computed in the following way

$$\begin{aligned} X_t &= (1+r)^{-(T-t)} \mathbb{E}_Q(X_T | \mathcal{F}_t) \\ &= (1+r)^{-(T-t)} \textstyle\sum_{j=0}^{T-t} \binom{T-t}{j} q^j (1-q)^{(T-t-j)} \varphi(u^j d^{(T-t-j)} s_0). \end{aligned} \tag{2.28}$$

Note 7: *An interesting feature of the binomial model proposed by Cox et al. (1979), which is in line with the one-period case, is that the natural probability P does not appear in the pricing functional, which is defined in terms of risk-neutral probability Q. This implies that several investors with various subjective probabilities will result in the same price.*

Additionally, a *replicating portfolio* or an *hedging portfolio* produces the same result as using the corresponding equivalent martingale measure approach. Since the binomial market model is complete, there exists a portfolio such that its value equals the value of the derivative at each time and at each node

$$V_t(j) = X_t(j), \tag{2.29}$$

where $j = 0, \dots, t$ is the number of "up" movements. The replicating portfolio at time t in node j is denoted by $\lambda_t^0(j)$ and $\lambda_t^1(j)$.

Proposition 2.1 (Bjork, 2009)
In a binomial market model, a derivative with maturity T and payoff $X_T = \varphi(S_T^1)$ can be *perfectly* replicated by a self-financing portfolio $\{\boldsymbol{\lambda}_t\}_{t=0}^{T-1}$. If $V_t(j)$ denotes the value of the replicating portfolio at time t in node with j "up" movements, with $j = 0, \dots, t$, then it can be recursively computed setting that

$$\begin{cases} V_t(j) = (1+r)^{-1}\left[qV_{t+1}(j+1) + (1-q)V_{t+1}(j)\right] \\ V_T(j) = X_T(j) = \varphi(u^j d^{T-j} s_0). \end{cases} \tag{2.30}$$

The replicating portfolio is given by $V_t(j) = \lambda_t^0(j)S_t^0 + \lambda_t^1(j)S_t^1(j)$, where

$$\begin{cases} \lambda_t^0(j) = \frac{1}{1+r}\frac{uV_t(j)-dV_t(j+1)}{u-d}, \\ \\ \lambda_t^1(j) = \frac{1}{S_{t-1}}\frac{V_t(j+1)-V_t(j)}{u-d}. \end{cases} \tag{2.31}$$

2.2.2 Trinomial model

The trinomial market model is the extension of the binomial one where the risky asset's price can evolve in three possible ways.

As the binomial, the trinomial market model is open at times $\{0, 1, \dots, T\}$, with $T \in \mathbb{N}$, and is composed by a riskless bond S^0 with price process as in (2.21), and by one risky asset S^1 with the following price process

$$S_0^1 = s_0 > 0, \quad S_t^1 = \begin{cases} uS_{t-1}^1 & \text{with probability } p_1 \\ mS_{t-1}^1 & \text{with probability } p_2 \\ dS_{t-1}^1 & \text{with probability } p_3 \end{cases} \tag{2.32}$$

for $t = 1, \dots, T$, where $u > m > d > 0$ are parameters, $p_i \in (0,1)$ for $i = 1, 2, 3$, and $\sum_{i=1}^3 p_i = 1$. The trinomial market model composed by only one risky asset is called *standard* trinomial market model.

As in the binomial framework, in order to have an equivalent martingale measure, we impose the martingality property

$$\mathbb{E}_Q\left(\frac{S_t^1}{S_{t-1}^1}|\mathcal{F}_0\right) = (1+r), \tag{2.33}$$

that leads to the following linear problem

$$\begin{cases} q_1 + q_2 + q_3 = 1, \\ q_i \in (0,1), & \text{for } i = 1,2,3 \ , \\ q_1 u + q_2 m + q_3 d = (1+r), \end{cases} \tag{2.34}$$

where $q_1 = Q(S_t^1 = uS_{t-1}^1)$, $q_2 = Q(S_t^1 = mS_{t-1}^1)$, and $q_3 = Q(S_t^1 = dS_{t-1}^1)$.

Problem (2.34) does not have, generally, a unique solution, then there exists a class of equivalent martingale measures denoted as

$$\mathcal{Q} = \left\{Q \in \mathbf{P}(\Omega, \mathcal{F}),\ Q \sim P :\ (1+r)^{-1}\mathbb{E}_Q(S_1^1) = S_0^1\right\}. \tag{2.35}$$

Hence, the *standard* trinomial model is incomplete.

We denote by $\mathrm{cl}(\mathcal{Q})$ and $\mathrm{ext}(\mathrm{cl}(\mathcal{Q}))$, respectively, the *closure* of $\mathcal{Q}$ and the set of *extreme points* of the closure of the class $\mathcal{Q}$.

The class $\mathcal{Q}$ is a convex set that can be characterized by the extreme points of its closure ($\mathcal{Q}$ is an open convex set, since it has to be composed by a set of strictly positive measures) (Runggaldier, 2006), that are

$$\begin{aligned}
Q^1 &= (q_1^{(1)}, q_2^{(1)}, q_3^{(1)}) &= \begin{cases} \left(0, \frac{(1+r)-d}{m-d}, \frac{m-(1+r)}{m-d}\right) & \text{if } m \geq (1+r), \\ \left(\frac{(1+r)-m}{u-m}, \frac{u-(1+r)}{u-m}, 0\right) & \text{if } m < (1+r), \end{cases} \\
Q^2 &= (q_1^{(2)}, q_2^{(2)}, q_3^{(2)}) &= \left(\frac{(1+r)-d}{u-d}, 0, \frac{u-(1+r)}{u-d}\right).
\end{aligned} \tag{2.36}$$

We stress that extreme points Q^1 and Q^2 are not equivalent to P since they are not strictly positive on $\mathcal{F}$; hence, the class of equivalent martingale measures is given by the strict convex combination of extreme points

$$\mathcal{Q} = \{Q^\alpha : Q^\alpha = \alpha Q^1 + (1-\alpha)Q^2, \alpha \in (0,1)\}. \tag{2.37}$$

Then $Q^\alpha \sim P$, for each $Q^\alpha \in \mathcal{Q}$. Since each $Q \in \mathcal{Q}$ is an equivalent martingale measure, it is consistent with the no-arbitrage assumption, but this can give rise to several prices for a derivative X with underlying asset S^1. An example of this problem is introduced in Example 2.1. in Cinfrignini (2022).

At this point we should define a suitable criterion to choose a specific $Q \in \mathcal{Q}$. Nevertheless, in this way we lose some information that are given by the whole class $\mathcal{Q}$. This problem is faced in the work of Cinfrignini et al. (2023) that will be presented in Chapter 4.

Anyhow, the incompleteness of the *standard* trinomial model can be straightforwardly overcome by considering the *completed* trinomial market model, that is composed by two risky assets, S^1 and S^2, each of them has price process as in (2.32) with parameters $u^1 \neq u^2$, $m^1 \neq m^2$ and $d^1 \neq d^2$. The martingale property (2.33) requires to solve the following linear problem

$$\begin{cases} q_1 + q_2 + q_3 = 1, \\ q_i \in (0,1), \\ u^1 q_1 + m^1 q_2 + d^1 q_3 = 1 + r, \\ u^2 q_1 + m^2 q_2 + d^2 q_3 = 1 + r. \end{cases} \tag{2.38}$$

It has a unique solution for q_1, q_2, q_3 and there exists a self-financing replication portfolio for every European derivative (for details, see Pascucci and Runggaldier, 2011).

We stress that each n-nomial model, with $n \geq 3$ and $(K+1)$ assets such that $(K+1) < n$, is *incomplete*. Nevertheless, each incomplete market model can be completed by adding $n-(K+1)$ risky assets, so that the class $\mathcal{Q}$ reduces to a singleton Q with a unique solution for $q_1, \ldots, q_n$.

Chapter 3

State of art

In this Chapter we review the main literature concerning frictional markets and *imprecise* stochastic processes.

In Section 3.1 we start with the literature on markets which do not satisfy the fundamental assumptions showed in Chapter 2: they are incomplete and/or they show the presence of frictions such as bid-ask spreads, taxation, transaction costs or other constraints on trading. We recall that an *incomplete* market model is characterized by a set of equivalent martingale measures and each derivative may not be perfectly replicated by a self-financing strategy.

Market incompleteness can be caused by a wide range of factors. It occurs when the number of (independent) risky assets is less than the market's sources of risk. One of the simplest example is the *standard* trinomial model. In this case, a technique to overcome the incompleteness is the *completion* of the market or the choice of a specific self-financing portfolio between the *imperfect* portfolios (that do not satisfy perfect replication or/and self-financing condition) such as the *super/sub-replicating portfolio*, or the choice of a specific equivalent martingale measure $Q \in \mathcal{Q}$ (we will see that the last two approaches are equivalent).

Another violation of the fundamental assumptions occurs when there are *frictions* or in presence of *ambiguity*, i.e., when we are not able to assign a specific model/probability measure that perfectly encodes the uncertainty of the market. They lead to the lack of *linearity* of the pricing rule. Then, literature focused on *non-linear* pricing rules, in particular defined in terms of the Choquet integral, thanks to the properties listed in Section 1.2.2.

Next, in Section 3.2, we review the main literature concerning *imprecise* random processes. Transition probabilities and/or the initial probability of a random process may be not precisely known or may come from different sources of uncertainty. In this setting, an approach could be to take the best estimates or choose one of the sources.

Another way to overcome the problem is to consider an *imprecise random process*, where the uncertainty measure, usually assumed to be a probability measure, is a non-additive measure. Usually, in this setting, one of the two properties between time-homogeneity and Markovianity can be lost. Moreover, the same generalization can be achieved from the expectation point of view, considering *non-linear expectations* such as the (coherent) lower expectation in Definition 1.7.

3.1 Incomplete and frictional markets

3.1.1 Replicating strategies for incomplete markets

A derivative traded on an incomplete market model, generally, cannot be replicated by a self-financing strategy. In the one-period setting, if $rank(A) \neq n \neq K+1$, Theorem 2.5 does not hold and the following possibilities and solutions can occur (Cerný, 2009):

- $rank(A) = n < (K+1)$. The market is *complete* since the number of assets is equal to the number of events but there are $K+1-n$ redundant assets that lead to $K+1-n$ free parameters in the solution of the problem $A\boldsymbol{\lambda}^T = \boldsymbol{X}$. In particular, a solution can be found by partitioning the matrix A into the matrix of n independent assets $A' \in \mathbb{R}^{n\times n}$ and the matrix of $(K+1-n)$ redundant assets $A'' \in \mathbb{R}^{n\times(K+1-n)}$. In turn, the vector of portfolio weights $\boldsymbol{\lambda}$ is partitioned into $\boldsymbol{\lambda}' \in \mathbb{R}^n$ and $\boldsymbol{\lambda}'' \in \mathbb{R}^{K+1-n}$ such that the linear problem can be equivalently written as

$$A\boldsymbol{\lambda}^T = A'(\boldsymbol{\lambda}')^T + A''(\boldsymbol{\lambda}'')^T = \boldsymbol{X}. \tag{3.1}$$

  Since redundant assets are linearly dependent, there exists a matrix C such that $A'' = A'C$; then (3.1) reduces to

$$A'(\boldsymbol{\lambda}')^T + A'C(\boldsymbol{\lambda}'')^T = \boldsymbol{X}. \tag{3.2}$$

  By arbitrarily choosing the portfolio of redundant weights $\boldsymbol{\lambda}''$, the portfolio of independent weights is computed by

$$\boldsymbol{\lambda}' = A_1^{-1}\boldsymbol{X} - C\boldsymbol{\lambda}''. \tag{3.3}$$

- $rank(A) = (K+1) < n$. There are no redundant assets but the market is *incomplete* since $n-(K+1)$ assets are lacking. The first method to complete the market is to decrease the number of states of the world, declaring that some states are "improbable", i.e., they have a zero probability (see, e.g., Melnikov, 1999). The second method, that is the most used, is to increase the number of assets. In particular, it is convenient to introduce $n-(K+1)$ linearly independent assets on the same probability space such that the completed system has a unique solution (e.g., Melnikov (1999) gives a criterion for a minimal completion of the market).

  Note 8: *Each completion of the matrix A, reached by adding the missing number of assets such that $rank(A) = n$, uniquely defines a single equivalent martingale measure in the class $\mathcal{Q}$ (Vasilev and Melnikov, 2021). In turn, working with all possible completions of the market is equivalent to work with the class $\mathcal{Q}$.*

- $rank(A) < n$, $rank(A) < (K+1)$. The market is *incomplete* and there are $(K+1) - rank(A)$ redundant assets. The market has to be completed and then, if $rank(A) < (K+1)$ occurs situation as in the first point.

Another technique that, unlike completion, does not alter the market structure is the selection of the "best" self-financing replicating strategy among the imperfect ones using approximations and algorithms. For instance, in the one-period setting, some criteria to do that are (Pascucci and Runggaldier, 2011):

sub(super)-hedging strategy. It looks for a strategy such that, in each $i \in \Omega$, the payoff of the replicating portfolio is at least less (greater) than the payoff of the derivative: $V_1^{\boldsymbol{\lambda}} \leq (\geq) X_1$. The sub-hedging price $\underline{V}_0^{\boldsymbol{\lambda}}$ and the super-hedging price $\overline{V}_0^{\boldsymbol{\lambda}}$ are the no-arbitrage bounds for the non-replicable payoff X_1.

quadratic risk minimization. It looks for a strategy that minimizes the expected value of the quadratic distance between the payoff of the derivative and the value of the portfolio. The following optimization problem has to be solved (for the application see, e.g., Bertsimas et al., 2001):

$$\min_{\boldsymbol{\lambda}} \mathbb{E}\left[(X_1 - V_1)^2\right].$$

shortfall risk minimization. It looks for a strategy that minimizes the shortfall risk. It penalizes only deviations in defect but it is less mathematically tractable. The following problem has to be solved:

$$\min_{\boldsymbol{\lambda}} \mathbb{E}\left[(X_1 - V_1)^+\right].$$

Equivalently, we can select an equivalent martingale measure in the class $\mathcal{Q}$ with a suitable criterion. For instance, Miyahara (1995) selects the probability measure that minimizes the relative entropy with respect to the natural probability.

Note 9: *Working with the envelopes of the set of equivalent martingale measures is equivalent to work with the supremum and infimum portfolios computed, respectively, in the set of all possible sub-replicating and super-replicating portfolios (see, e.g., Melnikov, 1999). In turn, choosing a specific $Q \in \mathcal{Q}$ is equivalent to choose a specific portfolio in the set of super(sub)-replicating portfolios.*

3.1.2 Frictional market models

Frictional markets have been examined in various ways in the literature depending on the goal of the study. The key literature on frictional markets in discrete time is examined in this section with a particular emphasis on frictions in the form of transaction costs and bid-ask spreads. We also examine the many methods suggested in literature for pricing financial contracts accounting for frictions.

First, we focus on the studies of Amihud and Mendelson (1986, 1991) which show the existence of frictions in the market (namely bid-ask spreads) and prove how they affect return and price of securities. Then, we take into account transaction costs and pricing strategies based on the replication of the derivative's terminal value, which can be *perfectly* replicated or *super*-replicated. Finally, we consider the bid-ask spread on prices and we review the literature regarding the existence and the properties of pricing rules allowing bid-ask spreads, focusing on pricing rule by means of the Choquet integral.

Empirical studies

Early research concentrated on determinants of bid-ask spread on prices. The term-to-maturity and yield-to-maturity have a positive relationship with bid-ask spread because a higher term-to-maturity or a higher yield-to-maturity means a higher risk of price change, according to Tanner and Kochin (1971) who examined Canada government bonds. Instead, a greater issue implies more buyer and sellers, which leads to a more liquid market, and a higher coupon rate indicates a reduced chance of price fluctuation. As a result, they have a negative relationship with bid-ask spread.

Amihud and Mendelson (1986) examined how bid-ask spreads affected returns of securities. They investigate the impact of liquidity on asset pricing under the assumption that the bid-ask spread is a measure of market liquidity. They believe that a security's expected return is an increasing and concave function of its bid-ask spread. This can be interpreted as follows: an investor facing a higher trading cost requires a higher return. However, there exists a *clientele effect*, which states that investors with longer holding periods hold stocks with wider spreads in order to cover transaction costs. As a result, returns of stocks with higher spread are less sensitive to the spread increase. They test these hypotheses on monthly returns and on the relative bid-ask spreads of NYSE stocks. They observe that their results support these assumptions. In turn, the influence of a firm-size variable, despite a negative association between company size and returns was seen (Banz, 1981; Reinganum, 1981a,b), is negligible once it was incorporated in their models.

Another research of Amihud and Mendelson (1991) looks at U.S. government securities with maturities under 6 months and tests the effect of liquidity on yield. They take Treasury Notes and Bills with same maturities[1]. These securities all share the same structure in terms of their tax, yield, maturity and duration, but their liquidity features vary. Security's liquidity is affected by the bid-ask spread and by any additional fees charged by brokers: smaller bid-ask spreads or fees are indicative of higher liquidity. Firstly, they examine at the distinct liquidity characteristics that quoted Bills and Notes exhibit, since they have, respectively, fees of 12.5\$ – 25\$ and 78\$ per million and a bid-ask spread of 1/128 of point and 1/32 of point. They make the assumption that Bills, since they have lower transaction costs, have a lower yield-to-maturity and they test this hypothesis on data. They show that the Notes' bid-ask spread is about 4 times greater that the Bills' spread and that the Notes' yield is greater than the Bills' yield. It proves how liquidity has a negative impact on yield-to-maturity and how yield-to-maturity is an inverse function of the liquidity.

Replicating strategies with transaction costs

One of the simplest methods that can be used to address the problem of pricing under the presence of transaction costs in a market model is setting up a replicating portfolio such that it replicates the value of a derivative in each node of the path and it takes into account transaction costs. Some authors have chosen to duplicate only

[1]Bills are short-term discount bonds whereas Notes are longer-term coupon-bearing bonds; however, Notes with maturity under 6 months become short-term single-payment securities like Bills.

the terminal value of the derivative, since a frictional model may be more challenging to implement.

As in the frictionless model, the replicating strategy under transaction costs is assumed to be self-financing and the no-arbitrage condition assures that the cost of the perfect replicating portfolio is the same as the cost of the derivative at the initial time. However, the replicating portfolio is not always the best option, since, as we will see throughout, there may be cheaper strategies that offer the same (or even bigger) payoff.

Garman and Ohlson (1980) develop a technique to assess the impact of *(non-proportional) transaction costs* in the equilibrium price of assets by avoiding arbitrage opportunities. They prove that there exists a linear operator such that the equilibrium price of assets coincides with the frictionless price added by an extra factor that they call *fudge factor*. Given a one-period market model defined over a finite filtered space $(\Omega, \mathcal{F})$, where, as usual, $\Omega = \{1, \ldots, n\}$ and $\mathcal{F}_T = \mathcal{F}_1 = \mathcal{P}(\Omega) = \mathcal{F}$, the market is composed by one asset with no transaction costs, with price process $\{S_0^0, S_1^0\}$ and by K risky assets with price processes $\{S_0^k, S_1^k\}$, for $k = 1, \ldots, K$, and with non-proportional state-contingent transaction costs at time $t = 0$ and $t = 1$ (i.e., the agent does not hold an initial position). There exist the following transaction costs:

- c_b^k, c_s^k are the transaction costs, respectively, to buy ($_b$) and sell ($_s$) the k-th asset at time $t = 0$;
- $c_b^k(i), c_s^k(i)$ are the transactions costs to buy ($_b$) and sell ($_s$) the k-th asset at time $t = 1$ in state $i \in \Omega$.

The no-arbitrage condition in presence of these transaction costs is defined as the non-existence of a portfolio $(\lambda_+^0, \lambda_-^0, \ldots, \lambda_+^K, \lambda_-^K)$, where $\lambda_+^k \geq 0$ is the share of the k-th asset bought and $\lambda_-^k \geq 0$ is the share of the k-th asset sold, with negative price and (at least) zero payoff. Formally, denoting by $(\lambda_+^k - \lambda_-^k) = \lambda^k$, a portfolio $(\lambda_+^0, \lambda_-^0, \ldots, \lambda_+^K, \lambda_-^K)$ does not allow an arbitrage opportunity if the following condition holds

$$\sum_{k=0}^{K} \lambda^k S_0^k + \sum_{k=0}^{K} (\lambda_+^k c_b^k + \lambda_-^k c_s^k) \geq 0 \Rightarrow \sum_{k=0}^{K} \lambda^k S_1^k(i) - \sum_{k=0}^{K} (\lambda_+^k c_s^k(i) + \lambda_-^k c_b^k(i)) \geq 0. \quad \text{(NA-GO)}$$

In the frictionless framework, the (NA-GO) condition is equivalent to the existence of a linear functional or a vector of state prices (i.e., we reduce to the classical results in Theorem 2.1). In turn, Garman and Ohlson (1980) want to reach the following representation of prices, analogous to the frictionless one

$$S_0^k = \sum_{i=1}^{n} D(i) S_1^k(i) + \epsilon^k, \tag{3.4}$$

where $\boldsymbol{D} = (D(1), \ldots, D(n)) \in \mathbb{R}^n$ is the vector of state price in the frictionless setting and ϵ^k is a component which directly depends on transaction costs, that they call *fudge factor*. They prove the following equivalence:

(GO.1) the market is (NA-GO);

(GO.2) there exists a non-negative state price vector $\boldsymbol{D} \in \mathbb{R}^n$ and K positive components $\{w^k, u^k\}$, with $k = 1, \dots, K$, such that (3.4) holds, where

$$\epsilon^k = c_s^k + \sum_{i=1}^{n} D(i)c_b^k(i) - u^k = -(c_b^k + \sum_{i=1}^{n} D(i)c_s^k(i) - w^k). \tag{3.5}$$

Despite the straightforward structure, Garman and Ohlson (1980) claim that transaction costs only affect assets' prices, while leaving the overall economy unchanged. It follows that the state price vector $\boldsymbol{D}$ can be computed in the classical frictionless model and (3.4) reduces to

$$S_0^k = \tilde{S}_0^k + \epsilon^k, \tag{3.6}$$

where $\tilde{S}_0^k$ is the price in the frictionless model.

From another point of view, Merton (1990) faces the problem of *proportional* transaction costs in a two-period binomial model in order to price a European call option through replicating strategy. The market is composed by one riskless asset (bond) S^0 without transaction costs and by one risky asset S^1 with proportional transaction costs. The bond's price process is

$$S_0^0 = 1,\ S_1^0 = 1 + r,\ S_2^0 = (1 + r)^2, \tag{3.7}$$

where $r > -1$ is the one-period risk-free interest rate. The risky asset's price evolves as in the binomial model

$$S_0^1 = s_0, \quad S_t^1 = \begin{cases} uS_{t-1}^1 \\ dS_{t-1}^1 \end{cases} \tag{3.8}$$

for $t = 1, 2$, with coefficients $u > d > 0$. The risky asset can be bought at its *ask* price ${}_aS_t^1 = (1 + \tau)S_t^1$ and sold at its *bid* price ${}_bS_t^1 = (1 - \tau)S_t^1$, where $\tau \geq 0$ is a fixed rate. As the frictionless binomial model, the Merton's model is free of arbitrage opportunities if and only if the following constraints are all satisfied:

$$us_0 < \quad (1 + r)s_0 \quad < \frac{ds_0(1 - \tau)}{1 + \tau}; \tag{3.9}$$

$$u^2 s_0 < \quad (1 + r)us_0 \quad < \frac{uds_0(1 - \tau)}{1 + \tau}; \tag{3.10}$$

$$uds_0 < \quad (1 + r)ds_0 \quad < \frac{d^2 s_0(1 - \tau)}{1 + \tau}. \tag{3.11}$$

Assuming that the agent does not have an initial position and all stocks held at maturity will be sold (this leads to assume the highest estimate of costs), Merton (1990) sets up a portfolio that perfectly replicates the payoff of a European call option C_T *at each node* of the path. The portfolio is denoted by $({}_\tau\boldsymbol{\lambda}_t) = ({}_\tau\lambda_t^0, {}_\tau\lambda_t^1)$ where ${}_\tau\lambda_t^0$ is the amount of S^0 held at time t after paid transactions costs for S^1, and ${}_\tau\lambda_t^1$ is the amount of S^1 held at time t after rebalancing. The price of the portfolio that undertakes a long position in a European call option with underlying asset S_2^1 is

$${}_\tau V_0 = V_0 + ({}_\tau\lambda_0^1 - \lambda_0^1)(s_0 - (1 + r)^{-1}us_0) + \tau {}_\tau\lambda_0^1(s_0 + (1 + r)^{-1}us_0), \tag{3.12}$$

where V_0 is the cost of the replicating portfolio in the frictionless setting and λ_0^1 is the weight on asset S^1 without transaction costs. For the computation of ${}_\tau\lambda_0^1$ we refer to Merton (1990) (equations (14.4a)-(14.4b)). ${}_\tau V_0$ is the call ask price, that is greater than its production cost with no transaction costs. Setting up a replicating strategy for the *short* position in the European call option, the author checks that it is not the reverse of the replicating strategy on the long position and he obtains the following relationship,

$$_{-\tau}V_0 < V_0 < {}_\tau V_0, \tag{3.13}$$

where ${}_{-\tau}V_0$ is the cost of the perfect replicating strategy for the short position in the same European call option. Finally, he stresses that in empirical examples the replicating portfolio's spread is larger than the τ spread in the underlying asset.

Boyle and Vorst (1992) extend the two-period model of Merton (1990) to *several periods*, setting up the replication portfolio at *each node* for a long and a short position in a European call option with *physical delivery settlement*. They show that if transaction costs tend to be zero, the model converges to the frictionless binomial model while, if the number of periods tend to be large, the closed formula they derive can be approximated by the formula of Black and Scholes (1973) with a modified variance. They set a self-financing replicating portfolio $\boldsymbol{\lambda}_t = (\lambda_t^0, \lambda_t^1)$ in each node of the process, assuming that the initial position $\boldsymbol{\lambda}_0$ was already hold by the agent (i.e., there are no transaction costs at initial time). In a two-period model, they point out that the self-financing portfolio is unique for the long call option if and only if

$$\lambda_2^1 \leq \lambda_0^1 \leq \lambda_1^1;$$

in turn, in the n-period model, the perfect replication of the short call option has a unique solution if and only if the following constraints are satisfied:

$$u(1-\tau) \geq (1+r)(1+\tau); \tag{3.14}$$

$$d(1+\tau) \leq (1+r)(1-\tau); \tag{3.15}$$

$$S_t^1 \notin \left[\tilde{K}(1+\tau)^{-1}; \tilde{K}(1-\tau){-1}\right], \tag{3.16}$$

where $\tilde{K}$ is the strike price of the call option[2]. Moreover, the long call price process, with proportional transaction costs τ, in a n-period model can be represented as a discounted expectation of an *adjusted process*. This adjusted process, denoted by $\{X_1, \ldots, X_n\}$, is a Markov process with two states and values $\ln u$ and $\ln d$ and with transition probability matrix given by

$$\overline{Q} = \begin{bmatrix} \overline{q}_u & \overline{q}_d \\ 1-\overline{q}_u & 1-\overline{q}_d \end{bmatrix}, \tag{3.17}$$

where

$$\overline{q}_u = \frac{(1+r)(1+\tau) - d(1-\tau)}{u(1+\tau) - d(1-\tau)}, \qquad \overline{q}_d = \frac{(1+r)(1-\tau) - d(1-\tau)}{u(1+\tau) - d(1-\tau)}.$$

[2]Palmer (2001) proves that, for a generic derivative that is not assumed to be physical delivery settled, the unique solution is reached just assuming that if ${}_\tau\lambda_1^1 <_\tau \lambda_2^1$ then $\frac{\tau(u+d)}{u-d} < 1$, otherwise $\tau < 1$.

The first column of $\overline{Q}$ represent the probability of X_{t+1} given that $X_t = \ln u$, the second column represents the probability of X_{t+1} given that $X_t = \ln d$. Hence, they characterize the following pricing formula

$$C_0 = (1+r)^{-n}\mathbb{E}_{\overline{Q}}\left[\left(\left(1+\overline{X}_n\tau\right)s_0 e^Y - \tilde{K}\right)\mathbf{1}_{\{s_0 e^Y \geq \tilde{K}\}}\right] \tag{3.18}$$

where $Y = \sum_{i=1}^n X_i$ and $\overline{X}_n = \begin{cases} 1 & \text{if } X_n = \ln u \\ -1 & \text{if } X_n = \ln d \end{cases}$.

If $\tau = 0$ the expected value reduces to the standard no-transaction costs pricing rule.

The authors conclude showing that, for large n and small transaction costs, the value of the replicating portfolio is approximately equal to the value computed with the Black-Scholes formula with a modified variance. For details, we refer to Theorem 3 of Boyle and Vorst (1992).

Bensaid et al. (1992) start from the model of Boyle and Vorst (1992) and they *relax the assumption of rebalancing at each node* of the process. They look for an optimal strategy among those dominating the payoff of the derivative at maturity and prove that, in some circumstances such as large transaction costs, the cost of a super-replicating strategy may be *lower* than the cost of the perfect replicating strategy as proposed by Boyle and Vorst (1992). This can be due to the fact that, when transaction costs are high, perfect replication is more expensive since it often requires to rebalance the portfolio. Given a binomial model with $T+1$ dates ($t = 0, \ldots, T$) defined on filtered state space ($\Omega = \{u, d\}^T, \mathcal{F}, \{\mathcal{F}_t\}_{t=0}^T$)[3], it is composed by a riskless asset S^0 and a risky asset S^1 whose price processes are $\mathcal{F}_t$-adapted. The riskless asset is constantly equal to 1, while the risky asset's price process follows a multiplicative binomial process with parameters $u > 1$ and $d < 1$, as in (3.8). There exist proportional transaction costs τ but they are not at the starting time neither at the expiration time: it means that the investor already has an initial position and at maturity the portfolio is not liquidated. For notation advantages, they define the transaction costs function

$$\psi(x) = \begin{cases} x(1+\tau) & \text{if } x \geq 0, \\ x(1+\tau)^{-1} & \text{if } x < 0. \end{cases} \tag{3.19}$$

As usual, a portfolio is a couple of adapted processes $\left(\{\lambda_t^0\}_{t=0}^{T-1}, \{\lambda_t^1\}_{t=0}^{T-1}\right)$ where $\lambda_t^0(i)$ denotes the amount of money (bond) holds in $[t, t+1]$ and $\lambda_t^1(i)$ denotes the share of stock S_t^1 holds in $[t, t+1]$, for $i \in \Omega$ and $t = 0, \ldots, T-1$. Given a derivative $X \in \mathbb{R}^\Omega$, they consider a weaker version of the self-financing condition at maturity, requiring that the portfolio holder can have an amount of cash greater than what is needed to replicate the derivative at maturity, then she/he has a greater payoff than the derivative's payoff, i.e., for all $i \in \Omega$, the following condition holds

$$\lambda_{T-1}^0(i) - \lambda_T^0(i) \geq S_T^1(i)\psi\left(\lambda_T^1(i) - \lambda_{T-1}^1(i)\right). \tag{SFw}$$

[3]We stress that in this framework, each $i \in \Omega$ is a particular path, i.e. a specific sequence of "up" and "down" movements.

They set up the following optimization problem $\mathcal{P}_1$:

$$\min_{\boldsymbol{\lambda}^0,\boldsymbol{\lambda}^1} \boldsymbol{\lambda}_0^0 + \boldsymbol{\lambda}_0^1 S_0^1,$$

subject to:

$$\begin{cases} \lambda_{T-1}^0(i) - \lambda_T^0(i) \geq S_T^1(i)\psi(\lambda_T^1(i) - \lambda_{T-1}^1(i)); \\ \lambda_{t-1}^0(i) - \lambda_t^0(i) = S_t^1(i)\psi(\lambda_t^1(i) - \lambda_{t-1}^1(i)), \quad \forall t \leq T-1 \text{ and } \forall i \in \Omega; \\ \lambda_{T-1}^0(i) + \lambda_{T-1}^1(i) S_T^1(i) = X_T(i). \end{cases} \tag{3.20}$$

The solution of problem $\mathcal{P}_1$ (denoted by $P_1(X)$) is the smallest cost of a portfolio with payoff at least equal to the payoff of the derivative. They prove that $P_1(X)$ always exists, it is bounded above by the perfect replication cost and it is *subadditive*, this means that $P_1(X) \geq -P_1(-X)$. Then, absence of arbitrage opportunities implies that

$$-P_1(-X) \leq \underline{X} \leq \overline{X} \leq P_1(X), \tag{3.21}$$

where $\underline{X}$ and $\overline{X}$ are, respectively, the ask price and the bid price of the derivative. They propose a recursive algorithm to compute the solution of $\mathcal{P}_1$ and they note that an agent does not trade at each date t if the portfolio's weights lies in a specific interval. It means that agents adjust their portfolio only in some circumstances, otherwise they continue to hold the position despite it does not perfectly replicate the derivative's value but, in this way, they save on transaction costs. They conclude by applying their algorithm to some kind of options and they obtain the following results: the perfect replicating strategy of Boyle and Vorst (1992) is the optimal solution for long European call option with physical delivery of the underlying asset and for both long and short calls when restrictions of Boyle and Vorst (1992) (3.14)–(3.16) are satisfied. Otherwise, for exotic call options and for both long and short call options with cash settlement, the perfect replicating strategy of Boyle and Vorst (1992) is *suboptimal* since the super-replicating strategy is cheaper.

Pliska (1997) stresses that the algorithm of Bensaid et al. (1992) can be difficult to apply and it is dependent on the assumption that options are automatically (not) exercised whenever $S_T^1(<) \geq \tilde{K}$, while there exist situations in which the option's holder would exercise or not the option depending on her/his preferences and portfolio composition. Then, Pliska (1997) reformulates the terminal condition in (3.20) and, relaxing the assumptions (3.14)–(3.16), presents an alternative method to get a closed-form solution in many cases.

Also, Edirsinghe et al. (1993) extend the model of Bensaid et al. (1992) considering *other constraints* such as size constraints and position limits. The market they propose shows the presence of different kinds of frictions: fixed transaction costs τ, variable transaction costs φ and quantity constraints, that is, the trading is restricted to be in multiples of a certain quantity δ_λ, such that

$$\lambda_t^1(i) - \lambda_{t-1}^1(j) = M\delta_\lambda, \tag{3.22}$$

where $M \in \mathbb{N}$.

They prove that, when transaction costs are low, the solution of their optimization problem is similar to that obtained by Boyle and Vorst (1992); on the contrary, when transaction costs arise, the replication cost proposed by Boyle and Vorst (1992) is larger than the minimum they obtain. This is due to the fact that Edirsinghe et al. (1993) require a super-replication condition also in the terminal constraint and it leaves the model free to reach a cheapest strategy. They conclude proposing a two-stage dynamic program that can solve the optimization problem and, differently from Boyle and Vorst (1992), it can be applied to non-convex payoffs.

Pricing rules with bid-ask spreads

The bid-ask spread is a kind of friction in a financial market that, differently from transaction costs we have just analysed, can be seen as an *implicit* transaction cost already incorporated by market prices. The *bid price* is the price such that a dealer offers to buy a financial contract while the a*sk price* is the price she/he offers to sell it. It follows that, in order to have a positive profit for the dealer, the ask price is greater that the bid price.

In this section, as usual, we denote a pricing rule as $\pi(\cdot)$, the bid price as $\underline{X}$ and the ask price as $\overline{X}$, for $X \in \mathbb{R}^\Omega$.

First of all, we point out that bid-ask price processes are usually assumed to be independent each others, then this study is different from the study of proportional transaction costs that, as we have seen (in particular in Boyle and Vorst, 1992), can be seen as a scaled price process.

Jouini and Kallal (1995a) study a one-period and a multi-period market model with bid-ask price processes for a derivative with *cash delivery settlement.* Given a filtered probability space $(\Omega, \mathcal{F}, P)$ (they do not require Ω to be finite), they start from a one-period model where $\mathcal{X} = L^2(\Omega, \mathcal{F}, P)$ denotes the set of all (square integrable) random variables and $\mathcal{X}_+$ the set of random variables $X \in \mathcal{X}$ such that $P(X \geq 0) = 1$ and $P(X > 0) > 0$ (i.e., there exists at least one $i \in \Omega$ such that $X(i) > 0$, in our formulation). They denote the set of positive linear functionals on $\mathcal{X}$ as Ψ, that is equivalent to consider a set of positive probability measures. The set of marketed claims is a convex cone[4] $\mathcal{M} \subseteq \mathcal{X}$ and each marketable contract $X \in \mathcal{M} \subseteq \mathcal{X}$ has a (upper) price at time $t = 0$ equal to $\pi(X)$, where $\pi : \mathcal{M} \to \mathbb{R}$ is a *sublinear* pricing rule[5].

They define a *free lunch* as *"a way to get a net payoff tomorrow arbitrarily close to a given positive claim at no cost today, or to get a net payoff tomorrow arbitrarily close to a non-negative claim for a negative cost today."* (Jouini and Kallal, 1995a, p.182).

Definition 3.1 (Free lunch) *Given a sequence of real numbers $r_n \in \mathbb{R}^n$, and two sequences of contingent claims $\{X_n\}, \{m_n\} \in \mathcal{M}$, a free lunch is given by the following conditions:*

(i) $\lim_{n\to+\infty} r_n = r^* \geq 0$;

[4] $\mathcal{M}$ is a convex cone of $\mathcal{X}$ if, for $X, Y \in \mathcal{M}$ and for all $\lambda \geq 0$ we have that $X + Y \in \mathcal{M}$ and $\lambda X \in \mathcal{M}$. We stress that this setting is a generalization in an *infinite* space of what we will set up in Section 4.3. In particular, as will be focused later, if we take a finite set $\Omega = \{1, \ldots, n\}$, $\mathcal{X} = \mathbb{R}^\Omega$ and a sublinear pricing rule π, approaches are comparable.

[5] Recall that sublinearity means that π satisfies both subadditivity and positively homogeneity and it is required to incorporate bid-ask spread since it leads to the following inequality: $\pi(X) \geq -\pi(-X)$.

(ii) $\lim_{n\to+\infty} X_n(i) = X^*(i) \geq 0$, *for all* $i \in \Omega$*;*

(iii) $r^* + X^*(i) \geq 0$ *for all* $i \in \Omega$ *with at least a strict inequality;*

(iv) $m_n(i) \geq X_n(i)$, *for all* $i \in \Omega$*;*

(v) $r_n + \pi(m_n) \leq 0$.

Theorem 3.1 (Jouini and Kallal, 1995a)
Given a one-period model with a price system $(\mathcal{M}, \pi)$, where $\mathcal{M}$ is a convex cone of $\mathcal{X}$ and π is a sublinear pricing rule on $\mathcal{M}$, the following conditions are equivalent:

(a) the price system $(\mathcal{M}, \pi)$ avoids free lunch;

(b) there exists a positive linear functional $\psi \in \Psi$, with $\psi : \mathcal{M} \to \mathbb{R}$, that lies under the sublinear pricing rule

$$\psi(X) \leq \pi(X),\ \forall X \in \mathcal{M}. \tag{3.23}$$

$\psi(\cdot)$ is called *underlying frictionless price* of $(\mathcal{X}, \pi)$ and, generally, it is not unique but each of them has to coincide on $\mathcal{M}$. The model reduces to the frictionless one with ψ and π that coincide, if π is linear.

Remark 2. Results of Jouini and Kallal (1995a) continue to hold in a finite state space $\Omega = \{1, \dots, n\}$, with $\mathcal{X} = \mathbb{R}^\Omega$ and $\mathcal{M} = \mathbb{R}^\Omega$.

Then, they extend the model to a multi-period setting with $t = 0, \dots, T$ with the *restriction* that the trading dates are a subset of $\{0, \dots, T\}$ and that they are defined in advance. The market is composed by K assets with ask price $\overline{S}_t^k$ and bid price $\underline{S}_t^k$, for $k = 1, \dots, K$, and by one ($k = 0$) frictionless asset constantly equal to 1, (i.e., amounts are already discounted), $\underline{S}_t^0 = \overline{S}_t^0 = 1$. This assumption does not reduce the generality of the model.

After requiring some technical assumptions[6], they prove that the absence of free lunch in the multi-period setting, in which consumption can occur only at initial and final date, is equivalent to the absence of free lunch in the induced one-period model $(\mathcal{M}, \pi)$, where $\mathcal{M}$ is the set of contracts that can be obtained by a self-financing strategy. The main theorem of Jouini and Kallal (1995a) is summarized as follows:

(JK.1) the frictional multi-period pricing model avoids multi-period free lunch[7] if and only if there exists at least an equivalent probability measure $Q \sim P$ and a process $\{S_t\}_{t=0}^T$ that lies between the bid and ask processes (i.e., $\underline{S}_t \leq S_t \leq \overline{S}_t$ for all t) such that $\{S_t\}_{t=0}^T$ is a martingale with respect to Q;

[6] They suppose that prices processes are *(i)* right-continuous and adapted to $\mathcal{F}_t$, *(ii)* $\mathbb{E}((\overline{S}_t^k)^2) < \infty$ and $\mathbb{E}((\underline{S}_t^k)^2) < \infty$ for all $k = 0, \dots, K$, and *(iii)* $\overline{S}_t^k > \underline{S}_t^k > 0$ for all t and for almost all $i \in \Omega$.

[7] A multi-period free-lunch is defined as a sequence of contingent claims $X_n \in \mathcal{M}$ converging to some $X^* \in \mathcal{X}_+$ such that there exists a self-financing strategy $(\lambda_t^+, \lambda_t^-)$ such that $(\lambda_t^+ - \lambda_t^-)^+ \overline{S}_T - (\lambda_t^+ - \lambda_t^-)^- \underline{S}_T \geq X_t$ for all t and $\lim_t \lambda_0^+ \underline{S}_0 - \lambda_0^- \overline{S}_0 \leq 0$.

(JK.2) there is a one-to-one correspondence between the set of linear functionals $\psi \in \Psi$ such that $\psi_{|M} \leq \pi$, and the set of expectation operators associated with Q:

$$Q(B) = \psi(\mathbf{1}_B), \quad \text{for all } B \in \mathcal{F}, \quad \text{and} \quad \psi(X) = \mathbb{E}_Q(X), \quad \text{for all } X \in \mathcal{X}; \tag{3.24}$$

(JK.3) for all marketable contracts $X \in \mathcal{M}$, we have that

$$[-\pi(-X), \pi(X)] = \text{cl}\{\mathbb{E}_Q(X), Q \sim P\}. \tag{3.25}$$

Point (JK.3) proves that the bid price of X is the smallest expected value with respect to martingale measures equivalent to P and, in turn, the ask price is the greatest expected value among the same set of equivalent martingale measures. Point (JK.1) shows that the no-arbitrage bid-ask price process can be seen as a *perturbation* of the frictionless no-arbitrage price process and that for each marketable contract[8], bid and ask prices are, respectively, the lower and the upper envelope of the set of expected values computed among each martingale measure equivalent to P.

In another research (see Jouini and Kallal, 1995b), they develop the same model as Jouini and Kallal (1995a) with a *sublinear* pricing rule and the set of contracts as a convex cone, adding more restrictions on the market: *short-selling constraints* and lending interest rate is different from the borrowing one. In a multi-period setting $t = 0, \ldots, T$, on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t=0}^T, P)$, they consider two kinds of securities: the first one cannot be short-sold while the second one can be *only* short-sold. They prove that this model is arbitrage-free if and only if there exists a positive process that acts as a numéraire and an equivalent probability measure such that the normalized (by the numéraire) price processes of securities that cannot be short-sold is a *super-martingale* and the normalized price processes of securities that can be just short-sold is a *sub-martingale*. Moreover, as in Jouini and Kallal (1995a), the arbitrage bounds on the bid-ask prices of a derivative are, respectively, the smallest and the largest expectation of the derivative's payoff with respect to all the numéraire processes and super-martingale probability measures. The interpretation they give is the following: when short-selling is prohibited, an asset can give the possibility of a loss without the possibility to cover this risk with a short-selling, then price processes only need to be non-increasing on average, i.e., super-martingales, to prevent arbitrage opportunities.

Roux (2011) studies a model under the same constraints of Jouini and Kallal (1995b) on a finite probability space $(\Omega, \mathcal{F}, P)$ with $\mathcal{F} = \mathcal{P}(\Omega)$ and P is strictly positive, with discrete time and finite time horizon $t = 0, \ldots, T$. The model is composed by K risky assets with proportional transaction costs, and the lending interest rate is different from the borrowing lending rate $r_t^B \geq r_t^L > -1$ for all t. He proves the following equivalence:

(R.1) there are no-arbitrage opportunities;

[8] A marketable contract is such that there exists a self-financing strategy composed by $K + 1$ assets that replicates the payoff of the contract X.

(R.2) there exists a probability measure $\tilde{P}$ (not necessarily equivalent to P), an adapted price process $\{S_t^*\}_{t=0}^T$ and an interest rate r^* such that, for all t,

$$\underline{S}_t^1 \leq S_t^* \leq \overline{S}_t^1, \quad r^L \leq r^* \leq r^B, \tag{3.26}$$

and the discounted price process $\{S_t^*/(1+r^*)\}_{t=0}^T$ is a martingale with respect to $\tilde{P}$.

Jouini (2000) extends the framework of Jouini and Kallal (1995a) to generic contracts $X \in \mathbb{R}^\Omega$ (here we take Ω to be finite), that can be not only cash settled, and he focuses on pricing rule properties. In particular, given the same filtered probability space in multi-period setting, composed by K assets with bid and ask prices and by one frictionless asset constantly equal to 1, such that the model satisfies the same assumptions of Jouini and Kallal (1995a) in footnote 6, they call a pricing rule $\pi : \mathbb{R}^\Omega \to \mathbb{R}$ *admissible* if the following axioms on the pricing rule hold:

(A1) it is sublinear;

(A2) it induces no-arbitrage, defined as the condition such that for each contract with $X \geq 0$, then $\pi(X) > 0$[9];

(A3) the price that π assigns to each contract has to be less or equal to the super-replication price obtained with respect to the smallest admissible portfolio, that is $\pi(X_T) \leq c(X_T)$, where $c(\cdot)$ is the smallest cost necessary to get a self-financing portfolio with at least the same payoff of X at time maturity;

(A4) it is lower semi-continuous.

The main results of Jouini (2000) are summarized as follows:

(J.1) there exists an *admissible* pricing rule π if and only if there exists (at least) an equivalent probability measure $Q \sim P$ and a process $\{S_t\}_{t=0}^T$ such that $\underline{S}_t^1 \leq S_t \leq \overline{S}_t^1$ for all t, and such that $\{S_t\}_{t=0}^T$ is a martingale with respect to Q;

(J.2) the *admissible* pricing functional is such that for all $X \in \mathbb{R}^\Omega$

$$\pi(X) \in \left[\inf_{Q \in \mathcal{Q}} \mathbb{E}_Q(X S_T^1), \sup_{Q \in \mathcal{Q}} \mathbb{E}_Q(X S_T^1) \right]. \tag{3.27}$$

Another stream of research is devoted to consider sets of probabilities in pricing, since this allows to model frictions in the form of bid-ask spreads. For a pricing theory to be accepted, a normative justification must be provided. For this, a generalization of the classical concept of arbitrage opportunity and the classical fundamental theorems of asset pricing are needed. In this vein, Carr et al. (2001) introduce a generalization of the classical arbitrage theory, in turn, generalizing the concept of arbitrage opportunity, in order to face the problem of pricing and hedging in incomplete markets. Given a market composed by a riskless bond with price

[9] The no-arbitrage condition of Jouini (2000) is *weaker* than the condition proposed by Jouini and Kallal (1995a) since the previous is defined in a static setting.

process $\{S_0^0 = 1, S_1^0 > 0\}$, and by K risky assets, with payoff $S_1^k(i)$, with $i \in \Omega$, and price S_0^k, they define an *opportunity* as a situation with price equal to zero. Agents specify a set of probability measures $\mathcal{P}$, indexed by $\mathcal{M} = \{1, 2, \dots, m\}$, that are divided into them (*stress probabilities*) such that, for every $X \in \mathbb{R}^\Omega$, the expected value is limited by a negative value

$$\mathbb{E}_{P^s}(X) \geq f_s, \tag{3.28}$$

where $f_s < 0$ and $s \in \mathcal{M}_s \subseteq \mathcal{M}$, and into them (*valuation probabilities*) such that the expected value is limited to zero

$$\mathbb{E}_{P^v}(X) = f_v = 0, \tag{3.29}$$

where $v \in \mathcal{M}_v \subseteq \mathcal{M}$ and $v + s = m$. In this setting, an *acceptable opportunity* is as an investment opportunity such that its expected value is greater than the floor of each possible probability (stress or valuation), i.e., it is weakly risky for a wide range of investors:

$$\mathbb{E}_{P^m}(X) \geq f_m, \quad \forall m \in \mathcal{M}. \tag{3.30}$$

An opportunity is *strictly acceptable* if probabilities are all *valuation measures*, i.e., $\mathcal{M}_v = \mathcal{M}$ and $\mathbb{E}_{P^m}(X) > 0$ for all $m \in \mathcal{M}$. The set of all strictly acceptable opportunities is denoted as $\mathcal{A}^+$. Given a vector of state prices $\boldsymbol{q}$, that is a vector such that $S_0^k = \frac{1}{S_1^0} \sum_{i \in \Omega} q(i) S_1^k(i)$, it is defined as a *representative state price vector* (RSPF) if there exists a vector of positive *"pricing weights"* $\boldsymbol{w}_v$, for all $v \in \mathcal{M}_v$, such that the state price vector is a combination of valuation probabilities, that is, for all $i \in \Omega$,

$$q(i) = \sum_{v \in \mathcal{M}_v} w_v P^v(i). \tag{3.31}$$

A strictly acceptable opportunity (SAO) does not exists if there is no portfolio $\boldsymbol{\lambda}$ such that

$$\sum_{k=1}^{K} \lambda^k S_0^k = 0, \quad \text{and} \quad \sum_{k=1}^{K} \lambda^k S_1^k \in \mathcal{A}^+. \tag{no-SAO}$$

They generalize the first fundamental theorem of asset pricing proving the equivalence between the (no-SAO) and the existence of RSPF. However, it does not prove the uniqueness of the RSPF. The generalization of completeness, called *acceptable completeness*, is defined as the existence of a portfolio $\boldsymbol{\lambda}$ such that, for all $m \in \mathcal{M}$,

$$\sum_{i \in \Omega} P^m(i) \sum_{k=1}^{K} \lambda^k S_1^k(i) - \sum_{i \in \Omega} P^m(i) X(i) = f_m. \tag{3.32}$$

The condition is equivalently satisfied if $K + 1 \geq |\mathcal{M}|$. We stress that this completeness condition is weaker than the classical one, hence, it introduces more opportunities. Under the assumption that $n \geq |\mathcal{M}|$, the generalization of the second fundamental theorem of asset pricing assures that the market is *acceptable complete* if and only if the RSPF is unique. It is equivalent to the condition that the number of assets $(K + 1) > v$. Moreover, if $v \leq K + 1 \leq n$, the market continues to be *acceptable complete* and then there exists a unique vector of *"pricing weights"*. They also generalize the definition of *ask* and *bid* price requiring, respectively, that they are the smallest value of the super-replicating portfolio that includes a possible loss of $f' \geq 0$ and the greatest value of the sub-replicating portfolio that includes the same possible loss of $f' \geq 0$. The pricing through the RSPF will lie between the bid and the ask prices.

Choquet pricing rules

A considerable piece of literature focused on a specific pricing functional that is expressed by the (discounted) Choquet integral. Thanks to its properties, the Choquet integral has a link with the standard expected value since it results to be the upper (lower) expectation among the expectations computed with respect to the probabilities in the core of a (at least) concave/convex capacity. This property, in a financial model that shows the presence of bid-ask spreads, supports the aim to obtain a *Choquet pricing rule* for the bid and the ask price such that they can be regarded as the upper and lower expectation of a set of probabilities.

One of the first contribution is given by Chateauneuf et al. (1996). They consider a one-period market which shows the presence of bid-ask spreads on prices, endowed with a pricing rule $\pi : \mathbb{R}^\Omega \to \mathbb{R}$ that assigns a (upper) price at time $t = 0$ to each payoff $X \in \mathbb{R}^\Omega$ available at time $t = 1$. The pricing rule satisfies the following properties:

(CKL.1) monotonicity: $X \geq Y \Rightarrow \pi(X) \geq \pi(Y)$, for all $X, Y \in \mathbb{R}^\Omega$;

(CKL.2) linearity on the frictionless asset: given the existence of $\mathbf{1}_\Omega = 1$ for all $i \in \Omega$, then $\pi(\alpha \mathbf{1}_\Omega) = \alpha$, for $\alpha \in \mathbb{R}$;

(CKL.3) subadditivity: $\pi(X + Y) \leq \pi(X) + \pi(Y)$, for all $X, Y \in \mathbb{R}^\Omega$ and with equality that holds only if X and Y are comonotone.

Let us note that (CKL.2) and (CKL.3) do not imply sublinearity since it should be required positive homogeneity on frictional assets and we also note that they consider already discounted amounts, i.e., $r = 0$, since the frictionless asset is constantly equal to 1.

The main result of Chateauneuf et al. (1996) is that if a pricing rule satisfies axioms (CKL.1)-(CKL.3) then there exists a unique *concave* capacity $\overline{\nu}$ such that the pricing rule is given by the Choquet expectation of the payoff

$$\pi(X) = \oint X \, \mathrm{d}\overline{\nu}. \tag{3.33}$$

By properties of the Choquet integral, the Choquet pricing rule is *sublinear* and it is the upper expectation with respect to the probability measures in the core($\overline{\nu}$) (see *(Ch.8)*, p.19). They suggest that the Choquet pricing rule can account for violations of the Put-Call parity relation that empirical researches proved to exist (see, e.g., Gould and Galai, 1974; Klemkosky and Resnick, 1979; Sternberg, 1984). Given a European put and call option with the same underlying asset X, strike price $\tilde{K}$ and expiration date (we suppose to be in a one-period setting, that is $t \in \{0, 1\}$), whose payoffs are $P_1 = \max(\tilde{K} - X_1, 0)$ and $C_1 = \max(X_1 - \tilde{K}, 0)$, the Put-Call parity relation is usually derived by the equality at maturity

$$P_1 = C_1 - X_1 + \tilde{K}. \tag{3.34}$$

Under a linear pricing rule, we should have

$$\pi(P_T) = \pi(C_T) - \pi(X_T) + \tilde{K}^{10}; \tag{CPP}$$

however, the non-linearity of the pricing rule (such as the Choquet pricing rule) explains the violations of the parity that are observed in real markets.

Bastianello et al. (2022) specify that if we require a pricing rule computed with respect to a *monotone* capacity to satisfy (CPP), it is a Choquet-Šipoš[11] pricing rule, denoted as $\pi_{CS}(\cdot)$, that does not allow bid-ask spreads since $\pi_{CS}(X) = -\pi_{CS}(-X)$. Chateauneuf and Cornet (2022b) remark the same result and they prove that if a Choquet pricing rule satisfies (CPP) and $\mathrm{core}(\overline{\nu}) \neq \emptyset$, that occurs if $\overline{\nu}$ is a concave capacity, then π is linear and ν is additive (i.e., it reduces to a probability measure).

In turn, Chateauneuf and Cornet (2022b) consider the following weaker version of (CPP):

$$\pi(P_1) \leq \pi(C_1) + \pi(-X_1) + \tilde{K}^{12}. \tag{CPP'}$$

They prove that, if a non-zero and *monotone* (they do not assume the subadditivity property) Choquet pricing rule satisfies (CPP'), it allows the presence of bid-ask spreads. Moreover, they prove that, given that π is a monotone Choquet pricing rule, the following conditions are equivalent:

(CC.1) $\mathrm{core}(\overline{\nu}) \neq \emptyset$ or equivalently $\overline{\nu}$ is a concave capacity (resp. $\mathrm{core}(\overline{\nu}) \neq \emptyset$ and it is strictly positive);

(CC.2) there exists a (resp. strictly positive) probability measure P such that $(1+r)^{-1}P \leq \overline{\nu}$;

(CC.3) π satisfies the following no-arbitrage condition (NA) (resp. (NA+)): $\forall X_k \in \mathbb{R}^\Omega$, with $k = 1, \ldots, K$, and for all K,

$$\sum_{k=1}^{K} X_k \geq 0 \Rightarrow \sum_{k=1}^{K} \pi(X_k) \geq 0. \tag{NA}$$

$$\left(\sum_{k=1}^{K} X_k > 0 \Rightarrow \sum_{k=1}^{K} \pi(X_k) > 0.\right) \tag{NA+}$$

[10] It is assumed that the risk-free interest rate of the market is $r = 0$, then the discounting factor is $(1 + r) = 1$; otherwise it should be $\tilde{K}(1 + r)^{-T}$.

[11] The Šipoš integral with respect to ν is $\int^S X \, \mathrm{d}\nu = \oint X^+ \, \mathrm{d}\nu - \oint X^- \, \mathrm{d}\nu$, for all $X \in \mathbb{R}^\Omega$. It does not allow bid-ask spread while the Choquet integral satisfies translation invariance. The Choquet and Šipoš integrals coincide if X is non-negative. Then, a pricing rule is a Choquet-Šipoš pricing rule if the Choquet integral and the Šipoš integral coincide when computed with respect to the same capacity.

[12] Chateauneuf and Cornet (2022b) suppose that there exists a riskless asset not constantly equal to 1, then the parity they show has the term $\tilde{K}\pi(\mathbf{1}_\Omega)$.

Condition (NA+) implies (NA) and if π is subadditive, the no-arbitrage conditions reduce to

$$\forall X \geq 0 \Rightarrow \pi(X) \geq 0, \tag{NAs}$$

$$\forall X > 0 \Rightarrow \pi(X) > 0. \tag{NAs+}$$

Assuming subadditivity of π, the no-arbitrage conditions (NAs) and (NA+) are satisfied, respectively, if and only if $\overline{\nu}$ is non-negative and $\overline{\nu}$ is strictly positive.

Results of Chateauneuf et al. (1996) and Chateauneuf and Cornet (2022b) can be summarized in this way: a pricing rule satisfying (CKL.1)-(CKL.3) (in particular subadditivity) is a Choquet pricing rule with respect to a concave capacity. It means that the core of the concave capacity is non-empty, then the Choquet pricing rule satisfies the no-arbitrage condition (NAs).

The relationship between put and call options prices is further studied in Cerreia-Vioglio et al. (2015) who axiomatically define the pricing rule. In a finite probability space $(\Omega, \mathcal{F}, P)$, they consider a one-period model with K risky assets, one riskless bond denoted as $\mathbf{1}_\Omega$, and a pricing rule $\pi : \mathbb{R}^\Omega \to \mathbb{R}$. They study the following put-call parity relation:

$$\pi(C_1) + \pi(-P_1) = \pi(X_1) - \tilde{K}\pi(\mathbf{1}_\Omega)^{13}. \tag{PCP}$$

Supposing that the market is complete, the main result of Cerreia-Vioglio et al. (2015) is given by the following equivalence:

(CV.1) a pricing rule π satisfies (PCP), it is monotone and positively homogeneous;

(CV.2) there exists a risk-neutral capacity $\overline{\nu}$ and a riskless interest rate $r > -1$ such that, for all $X \in \mathbb{R}^\Omega$, the pricing rule is given by the discounted Choquet integral

$$\pi(X) = (1+r)^{-1} \oint X \, \mathrm{d}\overline{\nu}. \tag{3.35}$$

If $\operatorname{core}(\overline{\nu}) \neq \emptyset$ (which occurs if $\overline{\nu}$ is concave), the pricing rule allows positive bid-ask spreads, since $\pi(X) \geq -\pi(-X)$. Chateauneuf and Cornet (2022b) stress that the converse is not true unless π is subadditive (see, e.g., Example 1 in Chateauneuf and Cornet, 2022b). Bastianello et al. (2022) go into detail of (PCP) and prove that the assumption of positively homogeneity in (CV.1) is redundant.

Additionally, at first Cerreia-Vioglio et al. (2015) do not require the subadditivity of the pricing rule as Chateauneuf et al. (1996) do, but they prove that assuming the pricing rule to be *subadditive*, the capacity is *concave* and the Choquet pricing rule agrees with the upper expectation with respect to its core. It follows that, thanks to the properties of concave capacities, its core is non-empty and it allows bid-ask spreads. In turn, condition (CC.1) of Chateauneuf and Cornet (2022b) is satisfied, then the subadditive Choquet pricing rule π assures the no-arbitrage condition (NA).

Bastianello et al. (2022) make a comparison between Put-Call parities studied by Chateauneuf et al. (1996) and Cerreia-Vioglio et al. (2015). First of all, we highlight that parities (PCP) and (CPP) are the same if π reduces to a linear

[13] We stress that Cerreia-Vioglio et al. (2015) do not consider that the riskless bond is constantly equal to 1; for this reason in (PCP) appears the term $\pi(\mathbf{1}_\Omega)$. However, it is a deterministic value since $\mathbf{1}_\Omega$ is riskless.

pricing rule, otherwise, we generally have that $\pi(-P) \neq -\pi(P)$. Without assuming particular properties on the pricing rule, Bastianello et al. (2022) prove the following equivalence:

(B.1) π satisfy parity (CPP);

(B.2) π satisfy parity (PCP) and there are no bid-ask spreads.

As already pointed out, the parity (CPP) is stronger and its achievement does not allow the presence of bid-ask spreads.

In this line of research, Chateauneuf and Cornet (2022a) consider a one-period model where assets' prices are characterized by the presence of bid-ask spread and they axiomatically define the *super-hedging* (or *super-replicating*) price of an arbitrage-free frictional market, denoted as $c(\cdot)$. It has to satisfy positively homogeneity, subadditivity, monotonicity and existence of a frictionless bond. Given a market composed by K frictional and risky assets with bid price $\underline{S}_0^k$ and ask price $\overline{S}_0^k$, for $k = 1, \dots, K$, and by a frictionless and riskless bond S^0 that it constantly equal to 1, i.e. $\underline{S}^0 = \overline{S}^0 = 1$, the market is *arbitrage free* (AF) if, for every portfolio $(\boldsymbol{\lambda}_+, \boldsymbol{\lambda}_-) \in \mathbb{R}^K \times \mathbb{R}^K$, where λ_+^k is the quantity of k-th asset bought and λ_-^k is the quantity of k-th asset sold, both conditions are verified:

$$\sum_{k=1}^{K} S_1^k(\lambda_+^k - \lambda_-^k) \geq 0 \Rightarrow \sum_{k=1}^{K} \lambda_+^k \overline{S}_0^k - \lambda_-^k \underline{S}_0^k \geq 0; \tag{PAF}$$

$$\sum_{k=1}^{K} S_1^k(\lambda_+^k - \lambda_-^k) > 0 \Rightarrow \sum_{k=1}^{K} \lambda_+^k \overline{S}_0^k - \lambda_-^k \underline{S}_0^k > 0. \tag{FAF}$$

For all $X \in \mathbb{R}^\Omega$, its *super-hedging price* is the cheapest portfolio with a payoff at least equal to X

$$c(X) = \inf_{(\boldsymbol{\lambda}_+, \boldsymbol{\lambda}_-) \in \mathbb{R}^K \times \mathbb{R}^k} \left\{ \sum_{k=1}^{K} \lambda_+^k \overline{S}_0^k - \lambda_-^k \underline{S}_0^k \ : \ \sum_{k=1}^{K} S_1^k(\lambda_+^k - \lambda_-^k) \geq X_1 \right\}. \tag{3.36}$$

The super-hedging price $c(\cdot)$ is a pricing rule since it satisfies the properties of finiteness, positively homogeneity, monotonicity, subadditivity and it allows the presence of a frictionless bond. The main result of Chateauneuf and Cornet (2022a) is that, given a market that is (PAF), the following assertions are equivalent:

(C.1) the market is *submodular*[14];

(C.2) the super-hedging pricing rule $c(\cdot)$ is a Choquet pricing rule with respect to a concave capacity.

[14] A market is submodular if its super-hedging price $c(\cdot)$ is such that $c(\max(X,Y)) + c(\min(X,Y)) \leq c(X) + c(Y)$ with $X, Y \in \mathbb{R}^\Omega$. For the financial interpretation, see Remark 2 in Chateauneuf and Cornet (2022a).

3.2 Imprecise stochastic processes

In real world, Markov and time-homogeneous processes may be difficult to characterize since transition probabilities as well as initial probabilities can be not precisely known because evidence supporting the occurrence of an event can be subjective or partial. Moreover, transition probabilities may be not constant in time (time-homogeneous). Then, a way to overcome this problem is to incorporate the imprecision into the model and relax the assumption of time-homogeneity, assuming that the initial distribution as well as transition probabilities are not precisely known and/or not constant in time.

Imprecise stochastic processes have been studied in literature starting from Hartfiel (1998) with the theory of *Markov set-chains* where uncertainty is introduced in parameters by means of intervals.

One of the first contribution in order to incorporate imprecision in a Markov process is given by Kozine and Utkin (2002) who construct *coherent interval-valued probabilities* in order to generalize discrete time Markov processes assuming that transition probabilities are time-homogeneous. As usual, we consider a finite filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t=0}^T, P)$, with $P(i) > 0$ for all $i \in \Omega$.

Kozine and Utkin (2002) start from the classical Markov and time-homogeneous process $\{X_t\}_{t=0}^T$ where

$$P(X_t = j) = \sum_{i=1}^n P(X_{t-1} = i)P(X_t = j|X_{t-1} = i) = \sum_{i=1}^n P(X_{t-1} = i)p_{i,j}, \quad (3.37)$$

for $t = 1, \dots, T$ and $i, j \in \Omega$.

They suppose that the initial probability α_i and the transition probabilities $p_{i,j}$ are not precisely known but they belong to intervals, denoted as

$$P(X_0 = j) = \alpha_j \in [\underline{\alpha}_j, \overline{\alpha}_j], \quad (3.38)$$

$$p_{i,j} \in [\underline{p}_{i,j}, \overline{p}_{i,j}], \quad (3.39)$$

such that bounds $\overline{\alpha}_j, \underline{\alpha}_j, \underline{p}_{i,j}$ and $\overline{p}_{i,j}$ are known, for all $i, j \in \Omega$.

The intervals of the state probability at time $t = 1$, denoted as $\underline{P}(X_1 = j)$ and $\overline{P}(X_1 = j)$, are computed by solving, for all $j \in \Omega$, the following problem

$$\underline{P}(X_1 = j) = \inf_{\substack{\alpha \in [\underline{\alpha}, \overline{\alpha}], \\ p_{i,j} \in [\underline{p}_{i,j}, \overline{p}_{i,j}]}} \sum_{i=1}^n \alpha_i p_{i,j}, \quad (3.40)$$

$$\overline{P}(X_1 = j) = \sup_{\substack{\alpha \in [\underline{\alpha}, \overline{\alpha}], \\ p_{i,j} \in [\underline{p}_{i,j}, \overline{p}_{i,j}]}} \sum_{i=1}^n \alpha_i p_{i,j}, \quad (3.41)$$

subject to:

$$\begin{cases} \sum_{i=1}^n \alpha_i = 1, & \\ \sum_{j=1}^n p_{i,j} = 1, & \text{for all } i \in \Omega, \\ \underline{\alpha}_i \leq \alpha_i \leq \overline{\alpha}_i, & \text{for all } i \in \Omega, \\ \underline{p}_{i,j} \leq p_{i,j} \leq \overline{p}_{i,j}, & \text{for all } i, j \in \Omega. \end{cases} \quad (3.42)$$

They prove that, given that the transition bounds $\underline{p}_{i,j}$ and $\overline{p}_{i,j}$ are coherent lower and upper probabilities (as defined in Section 1.2.1) and that initial bounds $\underline{\alpha}_i$, $\overline{\alpha}_i$ avoid sure loss[15], problem (3.40)–(3.42) reduces to the following linear programming problem

$$\underline{P}(X_t = j) = \inf_P \sum_{i=1}^{n} P(X_{t-1} = i)\underline{p}_{ij}, \tag{3.43}$$

$$\overline{P}(X_t = j) = \sup_P \sum_{i=1}^{n} P(X_{t-1} = i)\overline{p}_{i,j}, \tag{3.44}$$

for $j = 1, \ldots, n$, subject to:

$$\begin{cases} \sum_{i=1}^{n} P(X_{t-1} = i) = 1, & \text{for } i = 1, \ldots, n, \\ \underline{P}(X_{t-1} = i) \leq P(X_{t-1} = i) \leq \overline{P}(X_{t-1} = i), & \text{for } i = 1, \ldots, n. \end{cases} \tag{3.45}$$

Nevertheless, they do not provide a closed form for the solution of the linear programming.

Škulj (2006, 2009) generalizes the model proposed by Kozine and Utkin (2002) in two ways: omitting the time-homogeneity assumption on transition probabilities and allowing that all subsets belong to intervals, not only singletons $j \in \Omega$. In particular, the latter generalization means that, instead of working with *probability intervals (PRI)*[16], he works with *(partially determined) interval probabilities*, as defined by Weichselberger (2000).

Given a measurable space $(\Omega, \mathcal{F})$ (not necessarily finite) with $\mathcal{F} = 2^{\Omega}$, an *interval probability* is a couple of functions $[\underline{P}, \overline{P}]$ such that $\underline{P}$ and $\overline{P}$ are two set functions on $\mathcal{F}$ with $\underline{P} \leq \overline{P}$ and $\underline{P}(\Omega) = \overline{P}(\Omega) = 1$. To each interval probability $[\underline{P}, \overline{P}]$ there is an associate set $\mathcal{M}$ of additive probability measures that lie between $\underline{P}$ and $\overline{P}$, that we usually call *core*. The interval probabilities are called *F-field* if $\underline{P}(A) = \inf_{P \in \mathcal{M}} P(A)$ and $\overline{P}(A) = \sup_{P \in \mathcal{M}} P(A)$, for all $A \in \mathcal{F}$.

Remark 3. The F-field property is different from *coherence* in the Walley's sense since coherence allows finitely additive probabilities while Weichselberger model only allows σ-additive probabilities. In the case of finite probability spaces the two properties coincide and we refer to *(e)* in Definition 1.5. Moreover, as coherent lower probabilities, the relation $\overline{P}(A) = 1 - \underline{P}(A^C)$ holds and it implies that $\underline{P}$ and $\overline{P}$ are monotone.

The initial probability is fixed

$$P(X_0 = j) = \alpha_j \geq \underline{P}^{(0)}, \tag{3.46}$$

[15] They define the principle of *avoiding sure loss* as the conditions such that $\overline{P}(\Omega) \geq 1$, $\underline{P}(\Omega) \leq 1$, $0 \leq \underline{P}(A) \leq \overline{P}(A) \leq 1$, for all $A \in \mathcal{F}$ and $i \in \Omega$.

[16] A probability interval (PRI) is a special case of partially determined interval probabilities where the domain of the lower and upper bound is the set of all elementary events $i \in \Omega$. They are given by the interval $[\tilde{\underline{P}}, \tilde{\overline{P}}]$. As the (completely determined) probability intervals, they are called *F-field* if they are the lower and the upper envelope of the set of probability measures that lie between them.

for $j \in \Omega$, while $\underline{P}^{(0)}$ is the lower probability at time $t=0$ that is constant for all $j \in \Omega$. The transition Markov (*non* time-homogeneous) probabilities are given in terms of interval stochastic matrices

$$P(X_{t+1}=j|X_t=i)=p_{i,j}^{(t)} \geq \underline{p}_i, \tag{3.47}$$

where $\underline{p}_i$ is constant between j and t, and it is the lower bound of the following interval stochastic matrix

$$\mathrm{p}=\begin{bmatrix} \underline{p}_1 & \overline{p}_1 \\ \vdots & \vdots \\ \underline{p}_n & \overline{p}_n \end{bmatrix}, \tag{3.48}$$

where $\mathcal{M}(\mathrm{p})$ denotes the set of all stochastic matrices $(p_{i,j})$ such that $p_i \geq \underline{p}_i$. The aim of Škulj (2006, 2009) is to compute the probability distribution at next steps with a procedure analogous to that in the probabilistic setting, that, we recall, allows to compute, for $t=1$, $P(X_1=j)=\sum_{i=1}^n P(X_0=i)p_{i,j}^{(1)}$.

In the non-additive framework, since the initial probability belongs to a set $\mathcal{M}(\alpha)$ and transition probabilities belong to the set of stochastic matrices $p^{(t)} \in \mathcal{M}(\mathrm{p})$, the set of probability distributions at time t is the set of all possible initial probability distributions multiplied by all possible sequences of transition matrices

$$\mathcal{M}_t=\{\alpha \cdot p^{(1)} \cdot \ldots \cdot p^{(t)} \ : \ \alpha \in \mathcal{M}(\alpha),\ p^{(i)} \in \mathcal{M}(\mathrm{p}),\ \text{for } i=1,\ldots,t\}. \tag{3.49}$$

Sets $\mathcal{M}_t$ are generally not structured as interval probability measures but Hartfiel (1998) proves that if the initial set $\mathcal{M}(\alpha)$ is convex and the set of transition matrices $\mathcal{M}(\mathrm{p})$ is convex with separately specified rows[17], then $\mathcal{M}_t$ is a convex set of probabilities, for every $t \in \mathbb{N}$.

In order to compute the lower bound of $\mathcal{M}_t$, denoted as $\underline{P}^{(t)}(X)=\inf_{P\in\mathcal{M}_t} P(X)$, Škulj (2006) obtains the following result: assuming that the probability at time t is given by an interval of probabilities with bounds $\underline{P}_t, \overline{P}_t$, he computes the ext(cl(core($\underline{P}_t$))), denoted as P^{π_A} (see Section 1.2.1). It follows that

$$\underline{P}^{(t+1)}(A)=\sum_{i=1}^n P^{\pi_A}(i)\underline{p}_i(A). \tag{3.50}$$

In the following example we apply the proposed procedure.

Example 3.1 *Given $\Omega=\{1,2,3\}$ and $\mathcal{F}=2^\Omega \setminus \{\emptyset,\Omega\}$, with $|\mathcal{F}|=6$, suppose to have the following initial lower probability*[18] *$\underline{P}^{(0)}$ and lower transition probabilities $\underline{p}_1, \underline{p}_2, \underline{p}_3$ (we omit braces and commas to avoid cumbersome notation):*

$\mathcal{F}$	1	2	3	12	13	23
$\underline{P}^{(0)}$	0.1	0.3	0.4	0.5	0.6	0.7
$\underline{p}_1$	0.5	0.1	0.1	0.7	0.7	0.4
$\underline{p}_2$	0.1	0.4	0.3	0.6	0.5	0.8
$\underline{p}_3$	0.2	0.2	0.4	0.5	0.7	0.7

[17] Rows of a matrix $A \in \mathcal{A}$ are separately specified if the i-th row of A, denoted as a_i, and the i-th row of another matrix $A' \in \mathcal{A}$, denoted as a'_i are such that replacing a'_i with a_i, the resulting matrix also belongs to $\mathcal{A}$.

[18] We stress that whenever $|\Omega| \leq 3$, a lower probability is 2-monotone.

Let us compute the lower bound of the first step state probability $\underline{P}^{(1)}$. The permutations π_A such that $\underline{p}_{\pi_A(i)}(A) \geq \underline{p}_{\pi_A(i+1)}(A)$ are:

$$\begin{array}{llllll} \pi_1: & <1,2,3> & \pi_2: & <2,3,1> & \pi_3: & <3,2,1> \\ \pi_{12}: & <1,2,3> & \pi_{13}: & <1,3,2> & \pi_{23}: & <2,3,1> \end{array}$$

and the probability measure for each permutation is reported below

	1	2	3
P^{π_1}	0.1	0.4	0.5
P^{π_2}	0.3	0.3	0.4
P^{π_3}	0.3	0.3	0.4
$P^{\pi_{12}}$	0.1	0.4	0.5
$P^{\pi_{13}}$	0.1	0.4	0.5
$P^{\pi_{23}}$	0.3	0.3	0.4

It follows that, by (3.50)*, the lower probability for step* $t = 1$ *is*

	1	2	3	12	13	23
$\underline{P}^{(1)}$	0.19	0.23	0.28	0.56	0.62	0.64

♦

In another study (see Škulj, 2009) he considers 2-*monotone* lower probabilities and develop two simplified procedures in order to compute the lower bound of the state probability at time t. The *backward procedure* keeps the initial probability $\underline{P}^{(0)}$ as fixed and computes $\underline{P}^{(t)}(A) = \oint c_A^{(t)} \,\mathrm{d}\underline{P}^{(0)}$, with $c_A^{(t)} = \left(\underline{p}_1^{(t)}(A), \ldots, \underline{p}_n^{(t)}(A)\right)$, for all $A \in \mathcal{F} \setminus \{\emptyset, \Omega\}$. It gives the precise estimates with a constant integrating measure. The *forward procedure* computes the lower state probability at each step and keeps as constant the integrated function $\underline{P}^{(t)}(A) = \oint c_A \,\mathrm{d}\underline{P}^{(t-1)}$ as constant, with $c_A = \left(\underline{p}_1(A), \ldots, \underline{p}_n(A)\right)$, for all $A \in \mathcal{F} \setminus \{\emptyset, \Omega\}$.

From another starting point, Kast et al. (2014) study an imprecise random process whose uncertainty is modeled by a capacity ν such that it allows the presence of ambiguity in agent's preferences. They start from a binomial process endowed with a capacity and they show that it converges to a type of Brownian motion as limit of the Choquet binomial random process. Such process is proved to be *dynamically consistent*. Their approach is *axiomatic* and *subjective*, since they do not refer to an objective probability measure that is *"influenced"* by subjective preferences. Given a symmetric binomial additive process, denoted as $\{S_t\}_{0 \leq t \leq T}$, with

$$S_t = \begin{cases} S_t(u) = S_{t-1} + 1 \text{ if "up"}, \\ S_t(d) = S_{t-1} - 1 \text{ if "down"}, \end{cases} \tag{3.51}$$

each $\{S_0, \ldots, S_t\}$ is a trajectory of the process and at each node $S_t(i)$, with $0 \leq t \leq T$ and $i = \{"u", "d"\}$, the conditional capacity is constant and it is

$$\nu(S_t(u)|S_{t-1}) = \nu(S_t(d)|S_{t-1}) = c, \tag{3.52}$$

with $0 < c < 1$. If $c = \frac{1}{2}$, ν reduces to a probability measure.

Assuming a set of axioms on preferences (see Appendix 1 in Kast et al., 2014), for any random variable $X \in \mathbb{R}^{\Omega}$, for all $\tau \in \{0, \ldots, T\}$ and for all $\mathcal{F}_\tau \subseteq \mathcal{F}$, the following condition holds:

$$\mathbb{C}_0(X_t) = \mathbb{C}_0 \left(\mathbb{C}(X_t | \mathcal{F}_\tau) \right). \tag{3.53}$$

A random process satisfying (3.53) is called *dynamically consistent Choquet random walk* and it is completely characterized by a unique capacity such that

$$\nu(S_t(u)|S_{t-1}) = \nu(S_t(d)|S_{t-1}) = c. \tag{3.54}$$

They stress that a binomial tree that is *path-dependent* cannot be determined by a unique capacity as done for a *path-independent* process since the dynamically consistency condition does not work because it would lead to different values at time t from the following "up" and "down" nodes. On the contrary, if $\{X_t\}_{0 \leq t \leq T}$ is a *symmetric* Choquet random walk, the Choquet expectation reduces to $\mathbb{C}(X_t) = t(2c - 1)$. Finally, they prove that, when the time interval converges to 0, the dynamically consistent Choquet random walk in (3.53) converges to a general Wiener process with mean $m = 2c - 1$ and variance $s^2 = 4c(1 - c)$.

Another field of research studied stochastic processes under *non-linear* expectations. For instance, Nendel (2021) gives a definition of *convex* Markov chain (see Definition 2.2 in Nendel, 2021) with respect to a *convex* non-linear expectation, defined (in a finite formulation) as $\tilde{\mathbb{E}} : \mathbb{R}^{\Omega} \to \mathbb{R}$ such that:

(i) $\tilde{\mathbb{E}}(X) \leq \tilde{\mathbb{E}}(Y)$, for all $X(i) \leq Y(i)$, for all $i \in \Omega$;

(ii) $\tilde{\mathbb{E}}(\alpha \mathbf{1}_{\Omega}) = \alpha$, for all $\alpha \in \mathbb{R}$;

(iii) $\tilde{\mathbb{E}}(\alpha X + (1 - \alpha) Y) \leq \alpha \tilde{\mathbb{E}}(X) + (1 - \alpha) \tilde{\mathbb{E}}(Y)$, for all $X, Y \in \mathbb{R}^{\Omega}$ and $\alpha \in [0, 1]$.

We also mention the study of de Cooman et al. (2016) in the setting of *coherent lower (upper) expectations*, (Definition 1.7). They study an imprecise Markov chain that is not a collection of precise Markov chains since the Markov property is satisfied by the *set* of transition probabilities, while each element of the set may not be Markovian. Starting from a *local* model, that is a coherent lower prevision about the uncertainty of X_{t+1} after having observed $\{X_0 = x_0, \ldots, X_t = x_t\}$, they extend it to a *global* lower (upper) conditional expectation function considering the entire path such that it satisfies Markov and time-homogeneity properties. An analogous model is developed by T'Joens et al. (2021) starting from a *local* conditional upper expectation of X_{t+1} and extending it to a *global* conditional upper expectation as an extension of the local, without requiring the property of coherence but assuming another set of properties. Among the set of compatible global upper expectations, they require other conditions in order to select the global expectation that is most conservative. However they generally do not require the Markov property.

Chapter 4

A one-period n-nomial pricing model under Dempster-Shafer uncertainty

In the line of research on incomplete and frictional market models, which are described in Section 3.1, in this Chapter we present a work that has been published in Cinfrignini et al. (2023).

We start with the study of a one-period n-nomial market model, which is composed by a risk-free asset with deterministic price process and a risky asset whose price process follows a n-nomial multiplicative process.

Since we lose the market completeness, there exists an infinite class $\mathcal{Q}$ of equivalent martingale measures. Also, the linearity of the price functional is lost and it can be recovered by completing the market with extra securities or by choosing one of the equivalent martingale measures by some suitable criterion of choice. The incompleteness of the market in a n-nomial market model generates a form of "objective" ambiguity as one needs to deal with the class of equivalent martingale measures.

After having showed that the lower envelope $\underline{Q}$ of the class of equivalent martingale measures is a *belief function*, we try to derive a lower pricing rule from it as a (discounted) Choquet expectation, in a way to take care of bid-ask spreads. However, the closure of the set of equivalent martingale measures $\mathcal{Q}$ does not coincide, in general, with $\mathrm{core}(\underline{Q})$. To overcome this problem, we could think to use a suitable closed subset $\mathcal{Q}' \subseteq \overline{\mathcal{Q}}$ to define a *lower pricing rule* as a (discounted) lower expectation. Unfortunately, in Section 4.2 we show that in general this problem could not have a solution.

The belief function's framework allows to incorporate naturally frictions in the market, nevertheless, for a Choquet pricing rule to be acceptable the classical notion of arbitrage must be generalized. For that, in Section 4.3 we reformulate a general one-period pricing problem over a finite state space in the framework of belief functions. We provide a generalized one-step no-Dutch book condition and a generalized one-step no-arbitrage condition for a lower price assessment based on the *partially resolving uncertainty* principle proposed by Jaffray.

We prove that the generalized one-step no-Dutch book condition is necessary and sufficient for the existence of a belief function whose corresponding (discounted) Choquet expectation functional agrees with the lower price assessment. Next, we prove an analogue of the *first fundamental theorem of asset pricing* in the context of belief functions and in the one-period setting, showing that the generalized one-step no-arbitrage condition that we propose is equivalent to the existence of a *strictly positive* belief function whose corresponding discounted Choquet expectation functional agrees with the lower price assessment.

Concerning the original problem of deriving a lower pricing rule from the "risk-neutral" belief function $\underline{Q}$ arising in the n-nomial market model, in Section 4.4 we show that the discounted Choquet expectation with respect to $\underline{Q}$ does not satisfy the generalized no-Dutch book condition. Firstly, we take a specific $Q_0 \in \mathcal{Q}$ and we ϵ-contaminate it with the set of probabilities in $\mathrm{cl}(\mathcal{Q})$. Although being positive, its lower envelope induces a discounted Choquet expectation that continues to be not consistent with the lower price assessment. Hence, the idea is to ϵ-contaminate the reference probability Q_0 with a belief function $\widehat{Bel}$ that is an inner approximation of $\underline{Q}$ and such that it satisfies the generalized no-arbitrage principle. We propose a procedure for determining a belief function which inner approximates $\underline{Q}$, complying only with the lower price assessment of the stock. The task is achieved by minimizing a suitable distance, subject to a system of linear constraints, similarly to Miranda et al. (2021, 2022); Montes et al. (2018, 2019). In this way we get an *equivalent inner approximating one-step Choquet martingale belief function* $\widehat{Bel}_\epsilon$.

Finally, we show that if we further require $\widehat{Bel}$ to comply also with the upper price assessment (arriving to an inner approximating *strong* one-step Choquet martingale belief function), both $\widehat{Bel}$ and $\widehat{Bel}_\epsilon$ reduce to probability measures.

4.1 A n-nomial market model and its envelopes

We consider a one-period n-nomial market model, with times $t = 0$ and $t = 1$, composed by a risk-free asset (bond) and a risky asset (stock), as the binomial model in Section 2.2.1. The prices of the two securities are modeled by the processes $\{S_0^0, S_1^0\}$ and $\{S_0^1, S_1^1\}$ defined on the filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_0, \mathcal{F}_1\}, P)$, with $\Omega = \{1, \ldots, n\}$, $P(\{i\}) = p_i > 0$ for all $i \in \Omega$, $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and $\mathcal{F}_1 = \mathcal{F} = \mathcal{P}(\Omega)$.

We assume that $S_0^0 = 1$ and $S_0^1 = s_0 > 0$, while the prices at the end of the period satisfy

$$\frac{S_1^0}{S_0^0} = 1 + r, \quad \text{and} \quad \frac{S_1^1}{S_0^1} = \begin{cases} m_1, & \text{with probability } p_1, \\ m_2, & \text{with probability } p_2, \\ \vdots & \\ m_n, & \text{with probability } p_n, \end{cases} \tag{4.1}$$

where $m_1 > m_2 > \cdots > m_n > 0$ and $1 + r > 0$.

As usual, $Q \in \mathbf{P}(\Omega, \mathcal{F})$ is said to be *equivalent* to P, in symbol $Q \sim P$, when $P(A) = 0 \iff Q(A) = 0$, for all $A \in \mathcal{F}$. To avoid cumbersome notation, in what follows, every element $Q \in \mathbf{P}(\Omega, \mathcal{F})$ is identified with the vector $Q \equiv (q_1, \ldots, q_n)^T \in [0, 1]^n$, where $Q(\{i\}) = q_i$, for all $i \in \Omega$.

In this model, the set of *equivalent martingale measures* is defined as

$$\mathcal{Q} = \{Q \in \mathbf{P}(\Omega, \mathcal{F}) \ : \ (1+r)^{-1}\mathbb{E}_Q(S_1^1) = S_0^1, Q \sim P\}. \tag{4.2}$$

As is well-known, this set is not empty if $m_1 > 1+r > m_n$, moreover, $\mathcal{Q}$ is convex but generally not closed. In the particular case $n = 2$ this set reduces to the singleton $\mathcal{Q} = \{Q\}$ where $Q \equiv (q, 1-q)$ defined in (2.25), with $u = m_1$ and $d = m_2$.

In what follows we assume $n > 2$ and, as usual, we denote by $\mathrm{cl}(\mathcal{Q})$ and $\mathrm{ext}(\mathrm{cl}(\mathcal{Q}))$ the *closure* of $\mathcal{Q}$ and the set of *extreme points* of the closure.

We characterize the properties of the set of equivalent martingale measures $\mathcal{Q}$ because, despite each $Q \in \mathcal{Q}$ complies with the no-arbitrage assumption, we cannot account the whole set. Hence, our aim is to work with the lower envelope $\underline{Q}$ of $\mathrm{cl}(\mathcal{Q})$ defined, for every $A \in \mathcal{F}$, as

$$\underline{Q}(A) = \min_{Q \in \mathrm{cl}(\mathcal{Q})} Q(A). \tag{4.3}$$

We first provide a characterization of $\mathrm{ext}(\mathrm{cl}(\mathcal{Q}))$.

Theorem 4.1
For $n > 2$, if $m_{s-1} > 1+r \geq m_s$ for some $s \in \{2, \ldots, n\}$ and $1+r \neq m_n$, let $I = \{1, \ldots, s-1\}$ and $J = \{s, \ldots, n\}$, then

$$\mathrm{ext}(\mathrm{cl}(\mathcal{Q})) = \{Q_{i,j} \in \mathbf{P}(\Omega, \mathcal{F}) \ : \ (i,j) \in I \times J\},$$

where $Q_{i,j} \equiv (0, \ldots, 0, q_i, 0, \ldots, 0, q_j, 0, \ldots, 0)^T$, with

$$q_i = \frac{(1+r) - m_j}{m_i - m_j} \quad \text{and} \quad q_j = \frac{m_i - (1+r)}{m_i - m_j}.$$

Proof. We have that $Q \equiv (q_1, \ldots, q_n)^T$ is an element of $\mathrm{cl}(\mathcal{Q})$ if and only if it solves the system

$$\begin{cases} \sum_{k=1}^n q_k = 1, \\ \sum_{k=1}^n m_k q_k = 1+r, \\ q_k \geq 0, & \text{for } k = 1, \ldots, n. \end{cases}$$

It is immediate to see that the coefficient matrix associated with the first two equations has full rank, so it admits infinite solutions depending on $n-2$ real parameters. The set of such solutions is a closed subset of $\mathbb{R}^n$, while $\mathrm{cl}(\mathcal{Q})$ is the intersection of such set with the non-negative orthant.

As the rank is 2 (see, e.g., Faigle et al., 2002), the set $\mathrm{ext}(\mathrm{cl}(\mathcal{Q}))$ can be generated by selecting all the possible pairs of distinct indices $i, j \in \{1, \ldots, n\}$, and verifying if the vector $Q_{i,j}$, defined as in the statement of the theorem, is a solution of the above system. In turn, since $Q_{i,j}$ is a solution of the above system if and only if $i \in I$ and $j \in J$, the claim follows.

In particular, if $1+r > m_s$, we have $q_i, q_j \in (0,1)$ for every $i \in I$ and $j \in J$. On the other hand, if $1+r = m_s$, we have $q_i, q_j \in (0,1)$ for every $i \in I$ and $j \in J \setminus \{s\}$, while for $j = s$

$$q_i = \frac{(1+r) - m_s}{m_i - m_s} = \frac{m_s - m_s}{m_i - m_s} = 0 \quad \text{and} \quad q_j = q_s = \frac{m_i - m_s}{m_i - m_s} = 1,$$

thus $Q_{i,j}$ reduces to

$$Q_{i,j} = Q_s \equiv (0,\dots,0,q_s = 1,0,\dots,0)^T.$$

$\square$

Now we provide a characterization of the lower envelope $\underline{Q}$.

Theorem 4.2
For $n > 2$, if $m_{s-1} > 1+r \geq m_s$ for some $s \in \{2,\dots,n\}$ and $1+r \neq m_n$, let $I = \{1,\dots,s-1\}$ and $J = \{s,\dots,n\}$, then, for every $A \in \mathcal{F}$,

$$\underline{Q}(A) = \begin{cases} 1 & \text{if } A = \Omega, \\ \frac{(1+r)-m_{\underline{j}}}{m_1 - m_{\underline{j}}} & \text{if } 1+r \neq m_s \text{ and } I \subseteq A \neq \Omega, \\ & \text{or } 1+r = m_s \text{ and } I \cup \{s\} \subseteq A \neq \Omega, \\ \frac{m_{\overline{i}}-(1+r)}{m_{\overline{i}} - m_n} & \text{if } J \subseteq A \neq \Omega, \\ 0 & \text{otherwise}, \end{cases}$$

where $\underline{j} = \min\{j \in J : j \notin A\}$ and $\overline{i} = \max\{i \in I : i \notin A\}$.

Proof. We first prove the case $1+r \neq m_s$. We have that $\underline{Q}(A) = 1$ if and only if for all $(i,j) \in I \times J$, $\{i,j\} \subseteq A$, and this happens if and only if $A = \Omega$. Moreover, $\underline{Q}(A) = 0$ if and only if there exists $(i,j) \in I \times J, A \subseteq \{i,j\}^c$, and this happens if and only if $I \not\subseteq A$ and $J \not\subseteq A$.

For the remaining A's, two situations can occur: either *(a)* $I \subseteq A \neq \Omega$ or *(b)* $J \subseteq A \neq \Omega$.

(a). If $I \subseteq A \neq \Omega$, then

$$\begin{aligned} \underline{Q}(A) &= \min_{(i,j)\in I\times J} \left[\mathbf{1}_A(i)\frac{(1+r)-m_j}{m_i - m_j} + \mathbf{1}_A(j)\frac{m_i-(1+r)}{m_i - m_j} \right] \\ &= \min_{\substack{(i,j)\in I\times J \\ i\in A, j\notin A}} \frac{(1+r)-m_j}{m_i - m_j}. \end{aligned}$$

Suppose $i \in I$ and let $j \in J$ be such that $j \notin A$, with $m_1 > m_i > 1+r > m_j$. Since

$$\frac{(1+r)-m_j}{m_1 - m_j} - \frac{(1+r)-m_j}{m_i - m_j} = \frac{((1+r)-m_j)(m_i - m_1)}{(m_1 - m_j)(m_i - m_j)} < 0$$

we have $\frac{(1+r)-m_j}{m_1-m_j} < \frac{(1+r)-m_j}{m_i-m_j}$. Suppose $j, j' \in J$ are such that $j, j' \notin A$ with $m_1 > 1+r > m_j > m_{j'}$. Since

$$\frac{(1+r)-m_j}{m_1 - m_j} - \frac{(1+r)-m_{j'}}{m_1 - m_{j'}} = \frac{((1+r)-m_1)(m_j - m_{j'})}{(m_1 - m_j)(m_1 - m_{j'})} < 0$$

we have $\frac{(1+r)-m_j}{m_1-m_j} < \frac{(1+r)-m_{j'}}{m_1-m_{j'}}$. Hence, if $\underline{j}$ is the minimum element of J such that $\underline{j} \notin A$ we have that

$$\underline{Q}(A) = \min_{\substack{(i,j)\in I\times J \\ i\in A, j\notin A}} \frac{(1+r)-m_j}{m_i - m_j} = \frac{(1+r)-m_{\underline{j}}}{m_1 - m_{\underline{j}}}.$$

(b). If $J \subseteq A \neq \Omega$, then

$$\begin{aligned}\underline{Q}(A) &= \min_{(i,j)\in I\times J}\left[\mathbf{1}_A(i)\frac{(1+r)-m_j}{m_i-m_j}+\mathbf{1}_A(j)\frac{m_i-(1+r)}{m_i-m_j}\right]\\ &= \min_{\substack{(i,j)\in I\times J\\ i\notin A, j\in A}}\frac{m_i-(1+r)}{m_i-m_j}.\end{aligned}$$

Suppose $j \in J$ and let $i \in I$ be such that $i \notin A$ with $m_i > 1+r > m_j > m_n$. Since

$$\frac{m_i-(1+r)}{m_i-m_j}-\frac{m_i-(1+r)}{m_i-m_n}=\frac{(m_i-(1+r))(m_j-m_n)}{(m_i-m_j)(m_i-m_n)}>0$$

we have $\frac{m_i-(1+r)}{m_i-m_j} > \frac{m_i-(1+r)}{m_i-m_n}$. Suppose $i, i' \in I$ are such that $i, i' \notin A$ with $m_i > m_{i'} > 1+r > m_n$. Since

$$\frac{m_i-(1+r)}{m_i-m_n}-\frac{m_{i'}-(1+r)}{m_{i'}-m_n}=\frac{((1+r)-m_n)(m_i-m_{i'})}{(m_i-m_n)(m_{i'}-m_n)}>0$$

we have $\frac{m_i-(1+r)}{m_i-m_n} > \frac{m_{i'}-(1+r)}{m_{i'}-m_n}$. Hence, if $\bar{i}$ is the maximum element of I such that $\bar{i} \notin A$ we have that

$$\underline{Q}(A) = \min_{\substack{(i,j)\in I\times J\\ i\notin A, j\in A}}\frac{m_i-(1+r)}{m_i-m_j} = \frac{m_{\bar{i}}-(1+r)}{m_{\bar{i}}-m_n}.$$

Finally, we prove the case $1+r = m_s$. As before, we have that $\underline{Q}(A) = 1$ if and only if $A = \Omega$. Moreover, $\underline{Q}(A) = 0$ if and only if $I \cup \{s\} \nsubseteq A$ and $J \nsubseteq A$.

For the remaining A's, two situations can occur: either *(a')* $I \cup \{s\} \subseteq A \neq \Omega$ or *(b')* $J \subseteq A \neq \Omega$. Situation *(b')* coincides with *(b)*, thus it is proved in the same way.

(a') If $I \cup \{s\} \subseteq A \neq \Omega$, then proceeding as in the proof of *(a)* we have

$$\begin{aligned}\underline{Q}(A) &= \min_{(i,j)\in (I\cup\{s\})\times J}\left[\mathbf{1}_A(i)\frac{(1+r)-m_j}{m_i-m_j}+\mathbf{1}_A(j)\frac{m_i-(1+r)}{m_i-m_j}\right]\\ &= \min_{\substack{(i,j)\in (I\cup\{s\})\times J\\ i\in A, j\notin A}}\frac{(1+r)-m_j}{m_i-m_j} = \frac{(1+r)-m_{\underline{j}}}{m_1-m_{\underline{j}}}.\end{aligned}$$

□

In the next theorem we characterize the Möbius inverse of $\underline{Q}$.

Theorem 4.3
For $n > 2$, if $m_{s-1} > 1+r \geq m_s$ for some $s \in \{2,\dots,n\}$ and $1+r \neq m_n$, let $I = \{1,\dots,s-1\}$ and $J = \{s,\dots,n\}$. Let $\mu : \mathcal{F} \to \mathbb{R}$ be the Möbius inverse of $\underline{Q}$.

Then, for every $A \in \mathcal{F}$,

$$\mu(A) = \begin{cases} \frac{(1+r)-m_s}{m_1-m_s} & \text{if } 1+r \neq m_s \text{ and } A = I, \\ \frac{(1+r)-m_{k+1}}{m_1-m_{k+1}} - \frac{(1+r)-m_k}{m_1-m_k} & \text{if } A = \{1,\dots,k\} \text{ and } I \subset A \neq \Omega, \\ \frac{m_{s-1}-(1+r)}{m_{s-1}-m_n} & \text{if } A = J, \\ \frac{m_{k-1}-(1+r)}{m_{k-1}-m_n} - \frac{m_k-(1+r)}{m_k-m_n} & \text{if } A = \{k,\dots,n\} \text{ and } J \subset A \neq \Omega, \\ 0 & \text{otherwise.} \end{cases}$$

Proof. We first prove the case $1+r \neq m_s$, by considering all the possibilities for $A \in \mathcal{F}$.

(a). If $I \not\subseteq A$ and $J \not\subseteq A$, then $\mu(A) = 0$. Indeed, by Theorem 4.2 we have that $\underline{Q}(B) = 0$ for every $B \subseteq A$ and this implies $\mu(B) = 0$ for every $B \subseteq A$.

(b). If $A = I$, then by Theorem 4.2 the only $B \subseteq A$ with $\underline{Q}(B) \neq 0$ is $B = A = I$. Hence,

$$\mu(I) = \underline{Q}(I) = \frac{(1+r)-m_s}{m_1-m_s},$$

as s is the minimum element of J not in $A = I$.

(c). If $A = J$, then by Theorem 4.2 the only $B \subseteq A$ with $\underline{Q}(B) \neq 0$ is $B = A = J$. Hence,

$$\mu(J) = \underline{Q}(J) = \frac{m_{s-1}-(1+r)}{m_{s-1}-m_n},$$

as $s-1$ is the maximum element of I not in $A = J$.

(d). If $A \neq \{1,\dots,k\}$ and $I \subset A \neq \Omega$, then $\mu(A) = 0$. To see this, let $A = \{1,\dots,k\} \cup B$ with $I \subseteq \{1,\dots,k\} \neq \Omega$, $B \neq \emptyset$, and $B \cap \{1,\dots,k+1\} = \emptyset$. Since for all $E \subseteq A$ not containing I we have $\underline{Q}(E) = 0$, we can write

$$\mu(A) = \sum_{I \subseteq E \subseteq A} (-1)^{|A \setminus E|} \underline{Q}(E).$$

For every $s-1 \leq t \leq k$, if E contains $\{1,\dots,t\}$ but not $\{1,\dots,t+1\}$, it follows that $\underline{Q}(E) = \frac{(1+r)-m_{t+1}}{m_1-m_{t+1}}$ and all of such sets are of the form $\{1,\dots,t\} \cup C$ with $C \subseteq F$, where $F = \{t+2,\dots,k\} \cup B$ if $t+2 \leq k$ and $F = B$ otherwise. Moreover, we have

$$\begin{aligned} &\underline{Q}(\{1,\dots,t\} \cup F) - \sum_{\substack{D \subseteq F \\ |D| = |F|-1}} \underline{Q}(\{1,\dots,t\} \cup D) \\ &+ \sum_{\substack{D \subseteq F \\ |D| = |F|-2}} \underline{Q}(\{1,\dots,t\} \cup D) + \dots + (-1)^{|F|} \underline{Q}(\{1,\dots,t\}) = 0, \end{aligned}$$

since all terms are equal in absolute value and the number of positive terms is equal to that of negative terms. In turn, this implies that $\mu(A) = 0$.

(e). If $A \neq \{k,\dots,n\}$ and $J \subset A \neq \Omega$, then $\mu(A) = 0$. The proof of this claim is analogous to point *(d)*.

(f). If $A = \{1,\dots,k\}$ and $I \subset A \neq \Omega$, i.e., $s \leq k \leq n-1$, then, taking into account points *(a)*–*(e)*,

$$\underline{Q}(A) = \sum_{t=s-1}^{k} \mu(\{1,\dots,t\}) = \frac{(1+r)-m_{k+1}}{m_1-m_{k+1}}.$$

Hence, we have that

$$\mu(A) = \underline{Q}(A) - \sum_{t=s-1}^{k-1} \mu(\{1,\ldots,t\}) = \frac{(1+r) - m_{k+1}}{m_1 - m_{k+1}} - \frac{(1+r) - m_k}{m_1 - m_k}.$$

(g). If $A = \{k,\ldots,n\}$ and $J \subset A \neq \Omega$, i.e., with $2 \leq k \leq s-1$, then proceeding as in point *(f)* we get $\mu(A) = \frac{m_{k-1}-(1+r)}{m_{k-1}-m_n} - \frac{m_k-(1+r)}{m_k-m_n}$.

(h). If $A = \Omega$, then $\mu(\Omega) = 0$. Indeed, by points *(a)–(g)*, for every $A \in \mathcal{F} \setminus \{\Omega\}$, $\mu(A) \geq 0$ and, in particular, μ is strictly positive on the families

$$\begin{aligned} \mathcal{C}_1 &= \{\{1,\ldots,s-1\},\{1,\ldots,s\},\ldots,\{1,\ldots,n-1\}\}, \\ \mathcal{C}_2 &= \{\{s,\ldots,n\},\{s-1,\ldots,n\},\ldots,\{2,\ldots,n\}\}, \end{aligned}$$

while it is 0 otherwise. By the properties of the Möbius inverse, it must be $\sum_{A\in\mathcal{F}} \mu(A) = 1$, and since

$$\sum_{A\in\mathcal{F}\setminus\{\Omega\}} \mu(A) = \sum_{A\in\mathcal{C}_1} \mu(A) + \sum_{A\in\mathcal{C}_2} \mu(A) = \frac{(1+r)-m_n}{m_1-m_n} - \frac{m_1-(1+r)}{m_1-m_n} = 1,$$

it follows that $\mu(\Omega) = 0$.

Finally, we prove the case $1+r = m_s$. Proceeding as in points *(a)–(g)* by taking $I \cup \{s\}$ in place of I, it is possible to show that μ is strictly positive on the families $\mathcal{C}_2$ and

$$\mathcal{C}'_1 = \{\{1,\ldots,s\},\ldots,\{1,\ldots,n-1\}\},$$

while it is 0 on $\mathcal{F} \setminus (\{\Omega\} \cup \mathcal{C}'_1 \cup \mathcal{C}_2)$. Thus, in analogy to point *(h)*, since

$$\sum_{A\in\mathcal{F}\setminus\{\Omega\}} \mu(A) = \sum_{A\in\mathcal{C}'_1} \mu(A) + \sum_{A\in\mathcal{C}_2} \mu(A) = 1,$$

it follows that $\mu(\Omega) = 0$.

$\square$

The previous theorem implies that $\underline{Q}$ is completely monotone (i.e., a belief function) and it can be expressed as the strict convex combination of two necessity measures N_1, N_2 defined on $\mathcal{F}$, as stated in the following corollary.

Corollary 4.1 *The lower probability $\underline{Q}$ satisfies the following properties:*

(i) $\underline{Q}$ is completely monotone, that is, for every $k \geq 2$ and every $A_1,\ldots,A_k \in \mathcal{F}$, it holds that

$$\underline{Q}\left(\bigcup_{i=1}^{k} A_i\right) \geq \sum_{\emptyset\neq I\subseteq\{1,\ldots,k\}} (-1)^{|I|+1}\underline{Q}\left(\bigcap_{i\in I} A_i\right);$$

(ii) there exist two necessity measures $N_1, N_2 : \mathcal{F} \to [0,1]$ and $\alpha \in (0,1)$ such that, for every $A \in \mathcal{F}$,

$$\underline{Q}(A) = \alpha N_1(A) + (1-\alpha) N_2(A).$$

Proof. Statement *(i)* is an immediate consequence of Theorem 4.3 since $\mu(A) \geq 0$, for every $A \in \mathcal{F}$. For statement *(ii)*, we prove only the case $1 + r \neq m_s$ as the other case can be proved similarly. By Theorem 4.3, the focal elements of μ form two chains ordered by set inclusion:

$$\begin{aligned}\mathcal{C}_1 &= \{\{1,\dots,s-1\}, 1,\dots,s,\dots,1,\dots,n-1\},\\ \mathcal{C}_2 &= \{\{s,\dots,n\},\{s-1,\dots,n\},\dots,\{2,\dots,n\}\}.\end{aligned}$$

Let $\alpha = \sum_{A \in \mathcal{C}_1} \mu(A) = \frac{(1+r)-m_n}{m_1-m_n}$ and $(1-\alpha) = \sum_{A \in \mathcal{C}_2} \mu(A) = \frac{m_1-(1+r)}{m_1-m_n}$ and define μ_1, μ_2 on $\mathcal{F}$ setting, for every $A \in \mathcal{F}$,

$$\mu_1(A) = \begin{cases} \frac{\mu(A)}{\alpha} & \text{if } A \in \mathcal{C}_1,\\ 0 & \text{otherwise,}\end{cases} \quad \text{and} \quad \mu_2(A) = \begin{cases} \frac{\mu(A)}{1-\alpha} & \text{if } A \in \mathcal{C}_2,\\ 0 & \text{otherwise.}\end{cases}$$

A simple verification shows that μ_1, μ_2 are non-negative Möbius inverses with nested focal elements, so they induce, respectively, two necessity measures N_1, N_2 on $\mathcal{F}$. Then, by construction we have that, for every $A \in \mathcal{F}$,

$$\underline{Q}(A) = \alpha N_1(A) + (1-\alpha) N_2(A).$$

$\square$

The following example shows the computation of the lower envelope $\underline{Q}$ and its representation as a strict convex combination of two necessity measures.

Example 4.1 *Let* $\Omega = \{1,2,3,4\}$ *and* $m_1 = 4$, $m_2 = 2$, $m_3 = \frac{1}{2}$, $m_4 = \frac{1}{4}$, *and* $1 + r = 1$. *As usual, to avoid cumbersome notation, we denote events omitting braces and commas. In this case we have* $I = \{1,2\}$, $J = \{3,4\}$ *and* $\mathrm{ext}(\mathrm{cl}(\mathcal{Q})) = \{Q_{1,3}, Q_{1,4}, Q_{2,3}, Q_{2,4}\}$ *inducing the* $\underline{Q}$ *reported below*

$\mathcal{F}$	$\emptyset$	1	2	3	4	12	13	14	23	24	34	123	124	134	234	Ω
$Q_{1,3}$	0	$\frac{15}{105}$	0	$\frac{90}{105}$	0	$\frac{15}{105}$	1	$\frac{15}{105}$	$\frac{90}{105}$	0	$\frac{90}{105}$	1	$\frac{15}{105}$	1	$\frac{90}{105}$	1
$Q_{1,4}$	0	$\frac{21}{105}$	0	0	$\frac{84}{105}$	$\frac{21}{105}$	$\frac{21}{105}$	1	0	$\frac{84}{105}$	$\frac{84}{105}$	$\frac{21}{105}$	1	1	$\frac{84}{105}$	1
$Q_{2,3}$	0	0	$\frac{35}{105}$	$\frac{70}{105}$	0	$\frac{35}{105}$	$\frac{70}{105}$	0	1	$\frac{35}{105}$	$\frac{70}{105}$	1	$\frac{35}{105}$	$\frac{70}{105}$	1	1
$Q_{2,4}$	0	0	$\frac{45}{105}$	0	$\frac{60}{105}$	$\frac{45}{105}$	0	$\frac{60}{105}$	$\frac{45}{105}$	1	$\frac{60}{105}$	$\frac{45}{105}$	1	$\frac{60}{105}$	1	1
$\underline{Q}$	0	0	0	0	0	$\frac{15}{105}$	0	0	0	0	$\frac{60}{105}$	$\frac{21}{105}$	$\frac{15}{105}$	$\frac{60}{105}$	$\frac{84}{105}$	1
μ	0	0	0	0	0	$\frac{15}{105}$	0	0	0	0	$\frac{60}{105}$	$\frac{6}{105}$	0	0	$\frac{24}{105}$	0

Since $\mu(A) \geq 0$ *for every* $A \in \mathcal{F}$, *then* $\underline{Q}$ *is a belief function, and we have that* $\mathcal{C}_1 = \{12, 123\}$ *and* $\mathcal{C}_2 = \{34, 234\}$, *with*

$$\alpha = \mu(12) + \mu(123) = \frac{21}{105} \quad \textit{and} \quad \beta = 1 - \alpha = \mu(34) + \mu(234) = \frac{84}{105}.$$

The Möbius inverses μ_1, μ_2 *and the corresponding necessity measures* N_1, N_2 *are defined below*

$\mathcal{F}$	$\emptyset$	1	2	3	4	12	13	14	23	24	34	123	124	134	234	Ω
μ_1	0	0	0	0	0	$\frac{15}{21}$	0	0	0	0	0	$\frac{6}{21}$	0	0	0	0
N_1	0	0	0	0	0	$\frac{15}{21}$	0	0	0	0	0	1	$\frac{15}{21}$	0	0	1
μ_2	0	0	0	0	0	0	0	0	0	0	$\frac{60}{84}$	0	0	0	$\frac{24}{84}$	0
N_2	0	0	0	0	0	0	0	0	0	0	$\frac{60}{84}$	0	0	$\frac{60}{84}$	1	1
$\alpha N_1 + \beta N_2$	0	0	0	0	0	$\frac{15}{105}$	0	0	0	0	$\frac{60}{105}$	$\frac{21}{105}$	$\frac{15}{105}$	$\frac{60}{105}$	$\frac{84}{105}$	1

♦

4.2 A lower pricing rule

As pointed out in Section 3.1.2, real markets show the presence of frictions, mainly in the form of bid-ask spreads, that translate in the non-linearity of the price functional. Since the n-nomial model leads to a set of equivalent martingale measures $\mathcal{Q}$, we could think to allow frictions in the market by defining a *lower pricing rule* $\underline{\pi}$ as a discounted lower expectation.

As is well-known (see, e.g., Delbaen and Schachermayer, 2006), given a random variable $X \in \mathbb{R}^\Omega$ expressing the payoff at time $t = 1$ of a contract, its no-arbitrage price at time $t = 0$ can be computed relying on the set of equivalent martingale measures $\mathcal{Q}$, by computing

$$\pi_*(X) = \min_{Q \in \mathrm{cl}(\mathcal{Q})} (1+r)^{-1}\mathbb{E}_Q(X) \quad \text{and} \quad \pi^*(X) = \max_{Q \in \mathrm{cl}(\mathcal{Q})} (1+r)^{-1}\mathbb{E}_Q(X). \tag{4.4}$$

It holds that (see, e.g., Pliska, 1997; Cerný, 2009):

- if $\pi_*(X) = \pi^*(X)$, then their common value $\pi(X)$ is the no-arbitrage price at time $t = 0$ of X;
- if $\pi_*(X) < \pi^*(X)$, then the no-arbitrage price $\pi(X)$ at time $t = 0$ of X belongs to the open interval $(\pi_*(X), \pi^*(X))$.

If we have two contracts with payoffs $X, Y \in \mathbb{R}^\Omega$ at time $t = 1$, then their no-arbitrage price intervals are

$$(\pi_*(X), \pi^*(X)) \quad \text{and} \quad (\pi_*(Y), \pi^*(Y)),$$

nevertheless, we are not free to choose a value in one interval independently of the other, as shown in the following example.

Example 4.2 *Let* $\Omega = \{1, 2, 3\}$*,* $m_1 = 4$*,* $m_2 = 2$*,* $m_3 = \frac{1}{2}$*,* $1 + r = 1$ *and* $S_0^1 = 20$*. In this case we have* $I = \{1, 2\}$*,* $J = \{3\}$ *and* $\mathrm{ext}(\mathrm{cl}(\mathcal{Q})) = \{Q_{1,3}, Q_{2,3}\}$ *inducing the* $\underline{Q}$ *reported below*

$\mathcal{F}$	$\emptyset$	1	2	3	12	13	23	Ω
$Q_{1,3}$	0	$\frac{21}{105}$	0	$\frac{84}{105}$	$\frac{21}{105}$	1	$\frac{84}{105}$	1
$Q_{2,3}$	0	0	$\frac{45}{105}$	$\frac{60}{105}$	$\frac{45}{105}$	$\frac{60}{105}$	1	1
$\underline{Q}$	0	0	0	$\frac{60}{105}$	$\frac{21}{105}$	$\frac{60}{105}$	$\frac{84}{105}$	1
μ	0	0	0	$\frac{60}{105}$	$\frac{21}{105}$	0	$\frac{24}{105}$	0

Consider the following payoffs at time $t = 1$

Ω	1	2	3
X	20	10	10
Y	10	10	20

We have that

$$\begin{aligned}
\pi_*(X) &= \min\{\mathbb{E}_{Q_{1,3}}(X), \mathbb{E}_{Q_{2,3}}(X)\} = 10, \\
\pi^*(X) &= \max\{\mathbb{E}_{Q_{1,3}}(X), \mathbb{E}_{Q_{2,3}}(X)\} = 12, \\
\pi_*(Y) &= \min\{\mathbb{E}_{Q_{1,3}}(Y), \mathbb{E}_{Q_{2,3}}(Y)\} = \frac{110}{7} \approx 15.7, \\
\pi^*(Y) &= \max\{\mathbb{E}_{Q_{1,3}}(Y), \mathbb{E}_{Q_{2,3}}(Y)\} = 18,
\end{aligned}$$

so, we can consider the price assessment $\pi'(S_1^1) = 20$, $\pi'(X) = 11$, $\pi'(Y) = 17$. It holds that the partial price assessments $\{\pi'(S_1^1), \pi'(X)\}$ and $\{\pi'(S_1^1), \pi'(Y)\}$ are arbitrage-free, while the global price assessment $\{\pi'(S_1^1), \pi'(X), \pi'(Y)\}$ is not, as there is no $Q \in \mathcal{Q}$ such that $\pi'(S_1^1) = \mathbb{E}_Q(S_1^1)$, $\pi'(X) = \mathbb{E}_Q(X)$, $\pi'(Y) = \mathbb{E}_Q(Y)$.

♦

Since the lower envelope of the set $\mathcal{Q}$ is a belief function, the idea is to define a lower pricing rule by means of the Choquet integral with respect to $\underline{Q}$. However, as the following example shows, we have that, for $n > 2$, $\mathrm{cl}(\mathcal{Q}) \neq \mathrm{core}(\underline{Q})$ in general. Follows that the Choquet expectation with respect to $\underline{Q}$ is the lower expectation among $\mathrm{core}(\underline{Q})$, but it is not the lower expectation among the closure of the set of equivalent martingale measures $\mathrm{cl}(\mathcal{Q})$ appeared from the market model (see discussion in *(Ch.8)*, p.19).

Example 4.3 *Consider $\Omega = \{1, 2, 3, 4\}$, $m_1 = 4$, $m_2 = 2$, $m_3 = \frac{1}{2}$, $m_4 = \frac{1}{4}$, $1 + r = 1$, $\mathcal{Q}$ and $\underline{Q}$ of Example 4.1.*

A straightforward computation shows that $\mathrm{cl}(\mathcal{Q}) \neq \mathrm{core}(\underline{Q})$, *since (see* (1.26), *p.18), assuming $S_0^1 = s_0 > 0$,*

$$\mathbb{C}_{\underline{Q}}\left(\frac{S_1^1}{S_0^1}\right) = \frac{54}{105} < 1 = \min_{Q \in \mathrm{cl}(\mathcal{Q})} \mathbb{E}_Q\left(\frac{S_1^1}{S_0^1}\right).$$

Recall that $\mathrm{ext}(\mathrm{core}(\underline{Q})) = \{Q^\sigma : \sigma \in \Sigma\}$, *where Σ is the set of permutations of Ω (see* (1.20), *p.15). In particular, taking the permutation $\sigma = \langle 1, 2, 3, 4 \rangle$ and defining the probability measure*

$$\begin{aligned}
Q^\sigma &\equiv \left(\underline{Q}(1), \underline{Q}(12) - \underline{Q}(1), \underline{Q}(123) - \underline{Q}(12), \underline{Q}(1234) - \underline{Q}(123)\right)^T \\
&\equiv \left(0, \frac{15}{105}, \frac{6}{105}, \frac{84}{105}\right)^T
\end{aligned}$$

we have that $Q^\sigma \notin$ $\mathrm{cl}(\mathcal{Q})$ *which further proves that* $\mathrm{cl}(\mathcal{Q}) \subset \mathrm{core}(\underline{Q})$.

♦

Since the lower envelope of $\mathrm{cl}(\mathcal{Q})$ cannot be used to define a Choquet expectation functional that agrees with lower expectations, we could look for a suitable closed subset $\mathcal{Q}' \subseteq \mathcal{Q}$ to define a *lower pricing rule*. However, the choice of $\mathcal{Q}'$ is not free of issues since a reasonable criterion should be provided. The most natural way to get $\mathcal{Q}'$ is to consider a finite $\mathcal{G} \subset \mathbb{R}^\Omega$, and a *lower price assessment* $\underline{\pi}' : \mathcal{G} \to \mathbb{R}$. Here, the problem is to look for a closed $\mathcal{Q}' \subseteq \mathcal{Q}$ such that

$$\underline{\pi}'(X) = \min_{Q \in \mathcal{Q}'} (1+r)^{-1}\mathbb{E}_Q(X), \text{ for every } X \in \mathcal{G}.$$

A first (trivial) constraint for $\underline{\pi}'$ is that, for every $X \in \mathcal{G}$,

$$\pi_*(X) < \underline{\pi}'(X) < \pi^*(X), \text{ if } \pi_*(X) < \pi^*(X),$$

and $\underline{\pi}'(X) = \pi_*(X) = \pi^*(X)$ otherwise. However, this trivial constraint does not assure the existence of such a $\mathcal{Q}'$, as shown in the following example.

Example 4.4 *Let Ω, m_1, m_2, m_3, $1+r$, S_0^1, X and Y as in Example 4.2. Consider the lower price assessment $\underline{\pi}'(S_1^1) = 20$, $\underline{\pi}'(X) = 11$ and $\underline{\pi}'(Y) = 17$. We have that there is no closed subset $\mathcal{Q}' \subseteq \mathcal{Q}$ such that the corresponding discounted lower expectation functional agrees with $\underline{\pi}'$, in fact the following system*

$$\begin{cases} q_1 + q_2 + q_3 = 1, & \\ 4q_1 + 2q_2 + \frac{q_3}{4} = 1, & \\ 20q_1 + 10q_2 + 10q_3 = 11, & \\ 10q_1 + 10q_2 + 20q_3 \geq 17, & \\ q_k \geq 0, & k = 1, 2, 3, \end{cases}$$

is not compatible. Notice that the constraint related to $\underline{\pi}'(S_1^1) = 20$ is not reported since it is implied by the second equation.

We stress that, more generally, for the above assessment there is no closed subset $\mathcal{Q}'' \subseteq \mathbf{P}(\Omega, \mathcal{F})$ whose corresponding discounted lower expectation functional agrees with $\underline{\pi}'$. To see this, it is sufficient to consider the above system and relax the second constraint in a greater than or equal to constraint, as this result in an incompatible system.

♦

If we take the lower prices of securities in $\mathcal{G}$ as fixed, then the non-existence of a suitable closed $\mathcal{Q}' \subseteq \mathcal{Q}$ forces us to depart from the set $\mathcal{Q}$, in a way to derive a consistent discounted lower expectation. Thus, we should face a problem of correction of the set $\mathcal{Q}$ that necessarily introduces some imprecision with respect to $\mathcal{Q}$.

On the other hand, instead of looking for a closed $\mathcal{Q}' \subseteq \mathcal{Q}$, we could try to derive a lower pricing rule from the lower envelope $\underline{Q}$, which has been proved to be a belief function. The most natural way to get a lower pricing rule is to consider a discounted Choquet expectation derived from the "risk-neutral" belief function $\underline{Q}$.

The framework of belief functions allows to incorporate "naturally" frictions in the market, nevertheless, for such a lower pricing rule to be acceptable the classical notion of arbitrage must be generalized. This will be the aim of the next section. By considering only the lower envelope $\underline{Q}$ we forget of the set $\mathcal{Q}$ and actually work with $\mathrm{core}(\underline{Q})$, thus also in this case we introduce some imprecision with respect to $\mathcal{Q}$.

4.3 A one-step generalized no-arbitrage principle

In this section we still refer to a finite measurable space $(\Omega, \mathcal{F})$, with $\Omega = \{1, \ldots, n\}$ and $\mathcal{F} = \mathcal{P}(\Omega)$, with the filtration $\{\mathcal{F}_0, \mathcal{F}_1\}$ with $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and $\mathcal{F}_1 = \mathcal{F}$. We assume that the measurable space is endowed with a belief function Bel encoding the market beliefs.

Throughout this section we assume $Bel(A) > 0$, for every $A \in \mathcal{F} \setminus \{\emptyset\}$. Such a belief function Bel plays the same role of the "real-world" probability measure P in the classical formulation of a one-period market model. For this, Bel can be dubbed as "real-world" belief function.

We follow the probabilistic interpretation as pointed out in Section 2.1. Indeed, the belief function Bel gives rise to $\mathrm{ext}(\mathrm{core}(Bel))$, which is a finite set of "extreme" probabilistic opinions on the market, that can be associated with some reference agents. Such probabilistic opinions, though possibly different, are assumed to agree on states that are perceived as "unrealistic". Hence, the positivity of Bel generalizes the classical assumption on market beliefs as reference agents can have different probabilistic opinions, though it still requires some agreement among them on "realistic" states.

Note 10: *A more general formulation can be given referring to an arbitrary belief function Bel, still adopting the notion of equivalent belief function given in Definition 2. In turn, this requires to keep track of events with null belief, that will play a role in the formation of prices: this is a non-trivial generalization that could be the subject of future research. However, this generalization could be questionable from the financial point of view since we may have an asset with non-negative and non-null payoff with zero price.*

Definition 4.1 *Given two belief functions $Bel, \widehat{Bel}$ on $\mathcal{F}$, we say that $\widehat{Bel}$ is equivalent to Bel, in symbol $\widehat{Bel} \sim Bel$, if $Bel(A) = 0 \iff \widehat{Bel}(A) = 0$, for every $A \in \mathcal{F}$.*

Let us stress that, since Bel is positive on $\mathcal{F} \setminus \{\emptyset\}$, $\widehat{Bel} \sim Bel$ if and only if the same holds for $\widehat{Bel}$ and this happens if and only if its Möbius inverse $\widehat{\mu}$ is positive over the singletons.

We still consider a one-period market model related to times $t = 0$ and $t = 1$ where there is a risk-free bond assuring the return $1 + r > 0$. Such a bond has price $S_0^0 = 1$ at time $t = 0$ and payoff $S_1^0 = 1 + r$ at time $t = 1$. Here, the goal is to allow frictions in the market by considering, for a random variable $X \in \mathbb{R}^\Omega$, a lower price $\underline{\pi}(X)$ at time $t = 0$ and, if available, a corresponding upper price $\overline{\pi}(X)$, with $\underline{\pi}(X) \leq \overline{\pi}(X)$, to be interpreted as bid-ask prices. In the case of the risk-free bond we assume absence of frictions, meaning that the lower price coincides with the upper price $\underline{\pi}(S_1^0) = \overline{\pi}(S_1^0) = S_0^0$, thus we simply call it price.

We consider a finite non-empty collection of random variables

$$\mathcal{G} = \{S_1^1, \ldots, S_1^m\} \subset \mathbb{R}^\Omega, \tag{4.5}$$

expressing random payoffs at time $t = 1$ and a lower price assessment $\underline{\pi}' : \mathcal{G} \to \mathbb{R}$ related to time $t = 0$. In analogy with the classical formulation of no-arbitrage pricing, we do not require the risk-free bond to be part of $\mathcal{G}$ as it possesses a special role being used as numéraire.

Our aim is to determine a necessary and sufficient condition for the existence of a belief function $\widehat{Bel} \sim Bel$ such that, for $k = 1, \ldots, m$, it holds that

$$(1+r)^{-1}\mathbb{C}_{\widehat{Bel}}(S_1^k) = \underline{\pi}'(S_1^k).$$

By the positive homogeneity property of the Choquet integral (see *(Ch.3)*, p.19),

$$(1+r)^{-1}\mathbb{C}_{\widehat{Bel}}(S_1^k) = \mathbb{C}_{\widehat{Bel}}((1+r)^{-1}S_1^k),$$

thus we can consider the discounted payoff $\tilde{S}_1^k = (1+r)^{-1}S_1^k$, for $k = 1, \ldots, m$, and write

$$\mathbb{C}_{\widehat{Bel}}(\tilde{S}_1^k) = \underline{\pi}'(S_1^k). \tag{4.6}$$

Also in this case we have that $\tilde{S}_1^0 = \mathbf{1}_\Omega$.

Here we assume to know only the lower price of every S_1^k, for $k = 1, \ldots, m$. This is not restrictive since, if also the upper price assessment $\overline{\pi}' : \mathcal{G} \to \mathbb{R}$ is available, then the problem can be reformulated by considering

$$\mathcal{G}' = \{S_1^1, \ldots, S_1^m, -S_1^1, \ldots, -S_1^m\} \tag{4.7}$$

together with $\underline{\pi}'' : \mathcal{G}' \to \mathbb{R}$ such that, for $k = 1, \ldots, m$,

$$\underline{\pi}''(S_1^k) = \underline{\pi}'(S_1^k) \quad \text{and} \quad \underline{\pi}''(-S_1^k) = -\overline{\pi}'(S_1^k). \tag{4.8}$$

As usual, a *portfolio* is a vector $\boldsymbol{\lambda} = (\lambda^0, \ldots, \lambda^m)^T \in \mathbb{R}^{m+1}$ as in Definition 2.2, referred to contracts with random payoffs in $\mathcal{G}$.

Here, in contrast with the probabilistic setting, we assume the *partially resolving uncertainty* principle proposed by Jaffray (Jaffray, 1989) according to which the agent may only acquire the information that an event $B \neq \emptyset$ occurs, without knowing which is the true state of the world $i \in B$. Further, we assume that the agent is systematically pessimistic in his/her quantitative evaluations. As such, both in computing his/her (discounted) payoff related to a portfolio of securities and in the corresponding gain, the agent considers all non-impossible events in $B \in \mathcal{U} = \mathcal{F} \setminus \{\emptyset\}$ further, for every $X \in \mathbb{R}^\Omega$, he/she considers the corresponding $[X]^{\mathbf{L}} \in \mathbb{R}^{\mathcal{U}}$ built taking minima of X

$$[X]^{\mathbf{L}}(B) = \min_{i \in B} X(i). \tag{4.9}$$

This is in contrast with the principle of *completely resolving uncertainty* which is usually tacitly adopted and amounts in assuming that the agent will always acquire the information on the true state of the world $i \in \Omega$.

Working under partially resolving uncertainty, the final (discounted) payoff of the portfolio is the function $Z_{\boldsymbol{\lambda}} : \mathcal{U} \to \mathbb{R}$ defined, for every $B \in \mathcal{U}$, as

$$Z_{\boldsymbol{\lambda}}(B) = \lambda^0 + \sum_{k=1}^{m} \lambda^k [\tilde{S}_1^k]^{\mathbf{L}}(B), \tag{4.10}$$

while we interpret the quantity $\pi_{\boldsymbol{\lambda}} = \lambda^0 + \sum_{k=1}^m \lambda^k \underline{\pi}'(S_1^k)$ as the hypothetical price at time $t = 0$ of the portfolio that we would have if we were in a situation of completely resolving uncertainty. Hence, we can define the function $G_{\boldsymbol{\lambda}} : \mathcal{U} \to \mathbb{R}$ setting, for every $B \in \mathcal{U}$,

$$G_{\boldsymbol{\lambda}}(B) = Z_{\boldsymbol{\lambda}}(B) - \pi_{\boldsymbol{\lambda}} = \sum_{k=1}^{m} \lambda^k \left([\tilde{S}_1^k]^{\mathbf{L}}(B) - \underline{\pi}'(S_1^k)\right), \tag{4.11}$$

that can be interpreted as a random gain under partially resolving uncertainty.

Theorem 4.4
The following conditions are equivalent:

(i) there exists a belief function $\widehat{Bel}$ such that $\mathbb{C}_{\widehat{Bel}}(\tilde{S}_1^k) = \underline{\pi}'(S_1^k)$, for $k = 1, \ldots, m$;

(ii) for every $\boldsymbol{\lambda} = (\lambda^0, \ldots, \lambda^m)^T \in \mathbb{R}^{m+1}$ it holds that

$$\min_{B \in \mathcal{U}} G_{\boldsymbol{\lambda}}(B) \leq 0 \leq \max_{B \in \mathcal{U}} G_{\boldsymbol{\lambda}}(B).$$

Proof. The proof can be obtained applying Theorem 4.1 in Coletti et al. (2020), working with the dual capacity of $\widehat{Bel}$, which is a plausibility function. Here we provide a direct proof for the sake of completeness.

Fix an enumeration of $\mathcal{U} = \{B_1, \ldots, B_{2^n-1}\}$. Condition *(i)* is equivalent to the solvability of the following system

$$\begin{cases} \mathbf{Ax} = \mathbf{b}, \\ \mathbf{x} \geq \mathbf{0}, \end{cases}$$

where $\mathbf{x} = (\widehat{\mu}(B_1), \ldots, \widehat{\mu}(B_{2^n-1}))^T \in \mathbb{R}^{2^n-1}$ is an unknown column vector, $\mathbf{A} \in \mathbb{R}^{(m+1)\times(2^n-1)}$ is the coefficient matrix with

$$\mathbf{A} = \begin{pmatrix} \mathbf{1}_\Omega^{\mathbf{L}}(B_1) & \cdots & \mathbf{1}_\Omega^{\mathbf{L}}(B_{2^n-1}) \\ [\tilde{S}_1^1]^{\mathbf{L}}(B_1) & \cdots & [\tilde{S}_1^1]^{\mathbf{L}}(B_{2^n-1}) \\ \vdots & & \vdots \\ [\tilde{S}_1^m]^{\mathbf{L}}(B_1) & \cdots & [\tilde{S}_1^m]^{\mathbf{L}}(B_{2^n-1}) \end{pmatrix},$$

and $\mathbf{b} = (1, \underline{\pi}'(S_1^1), \ldots, \underline{\pi}'(S_1^m))^T \in \mathbb{R}^{m+1}$.

By Farkas' lemma (Theorem 1.3), the system above is compatible if and only if the following system is not compatible

$$\begin{cases} \mathbf{A}^T\mathbf{y} \leq \mathbf{0}, \\ \mathbf{b}^T\mathbf{y} > 0, \end{cases}$$

where $\mathbf{y} = (\lambda^0, \ldots, \lambda^m)^T \in \mathbb{R}^{m+1}$ is an unknown column vector. It holds that $\mathbf{A}^T\mathbf{y} \in \mathbb{R}^{2^n-1}$ and, for $i = 1, \ldots, 2^n - 1$, the ith component of constraint $\mathbf{A}^T\mathbf{y} \leq \mathbf{0}$ is

$$\lambda^0 + \sum_{k=1}^{m} \lambda^k [\tilde{S}_1^k]^{\mathbf{L}}(B_i) \leq 0,$$

moreover, subtracting the positive quantity $\mathbf{b}^T\mathbf{y}$ we get

$$\sum_{k=1}^{m} \lambda^k \left([\tilde{S}_1^k]^{\mathbf{L}}(B_i) - \underline{\pi}'(S_1^k)\right) < 0.$$

Thus, condition *(i)* is equivalent to the existence of $i \in \{1, \dots, 2^n - 1\}$ such that the above inequality does not hold, which, in turn, is equivalent to *(ii)*.

$\square$

The above theorem says that, working under partially resolving uncertainty, in order to have a discounted totally monotone Choquet expectation representation of the lower price assessment $\underline{\pi}'$ independent of the "real-world" belief function of the market, it is necessary and sufficient that every portfolio $\boldsymbol{\lambda}$ does not give rise to a sure loss or a sure gain over $\mathcal{U}$. In other terms, the above condition can be considered a *generalized (one-step) avoiding Dutch book condition*, working under partially resolving uncertainty.

Note 11: *Theorem 4.4 specializes results of Jouini and Kallal (1995a) summarized in Theorem 3.1 in a finite setting. Given that the generalized no-Dutch book condition (ii) holds, there exists a belief function $\widehat{Bel}$ that extends the price assessment to $\mathbb{R}^\Omega$ and gives rise to the lower pricing rule $\underline{\pi} : \mathbb{R}^\Omega \to \mathbb{R}$ such that $\underline{\pi}(\cdot) = (1+r)^{-1}\mathbb{C}_{\widehat{Bel}}(\cdot)$. Each belief function $\widehat{Bel}$ defines its conjugate plausibility function $\widehat{Pl}$, with respect to which we define an upper pricing rule $\overline{\pi}(X) = (1+r)^{-1}\mathbb{C}_{\widehat{Pl}}(X)$, for all $X \in \mathbb{R}^\Omega$. Considering that $\mathcal{M}$ is the convex cone generated by $\mathcal{G}$, and $\mathcal{X} = \mathbb{R}^\Omega$, then $\overline{\pi}'$ gives rise to a restriction of $\overline{\pi}$ on $\mathcal{M}$, that is analogue to the (upper) pricing rule considered by Jouini and Kallal (1995a). Each belief/plausibility function implies the existence of a non-empty $\mathrm{core}(\widehat{Bel}) \equiv \mathrm{core}(\widehat{Pl}) = \{P \in \mathbf{P}(\Omega, \mathcal{F}) \ : \ P \geq \widehat{Bel}\} \neq \emptyset$, that is equivalent to say that there exists (at least) a probability measure P such that the expectation with respect to that is dominated by the upper expectation*

$$(1+r)^{-1}\mathbb{E}_Q(X) \leq (1+r)^{-1}\mathbb{C}_{\widehat{Pl}}(X) = \overline{\pi}'(X), \tag{4.12}$$

for all $X \in \mathcal{M}$.

The functional $\overline{\pi}'$ is sublinear over $\mathcal{M}$ thanks to the properties of the Choquet integral, then condition (b) of Theorem 3.1 is satisfied. It follows that our $\widehat{Bel}$ satisfies the no free-lunch condition in Definition 3.1.

Nevertheless, the condition *(ii)* of Theorem 4.4 does not assure that $\widehat{Bel} \sim Bel$, that is we do not have any guarantee that $\widehat{Bel}(A) > 0$, for every $A \in \mathcal{U}$.

We stress that, like in the classical no-arbitrage theory (see Note 3), the request of positivity of $\widehat{Bel}$ is a desideratum in the context of pricing since it assures that a security with a non-negative and non-null payoff at time $t = 1$ will have a positive lower price at time $t = 0$. Hence, we provide a necessary and sufficient condition for the existence of an equivalent belief function positive on the entire $\mathcal{U}$ that can be dubbed "risk-neutral" belief function. Besides equivalence between $\widehat{Bel}$ and Bel no other relation is asked to hold.

Such theorem is the analogue of the *first fundamental theorem of asset pricing*, formulated in the Dempster-Shafer theory of evidence.

Theorem 4.5
The following conditions are equivalent:

(i) there exists a belief function $\widehat{Bel} \sim Bel$, i.e., $\widehat{Bel}(A) > 0$, for every $A \in \mathcal{U}$, such that $\mathbb{C}_{\widehat{Bel}}(\tilde{S}_1^k) = \underline{\pi}'(S_1^k)$, for $k = 1, \dots, m$;

(ii) for every $\boldsymbol{\lambda} = (\lambda^0, \dots, \lambda^m)^T \in \mathbb{R}^{m+1}$ none of the following conditions holds:

(a) $Z_{\boldsymbol{\lambda}}(\{i\}) = 0$, for $i = 1, \dots, n$, $Z_{\boldsymbol{\lambda}}(B) \geq 0$, for all $B \in \mathcal{U} \setminus \{\{i\} : i \in \Omega\}$ and $\pi_{\boldsymbol{\lambda}} < 0$;

(b) $Z_{\boldsymbol{\lambda}}(\{i\}) \geq 0$, for $i = 1, \dots, n$, with at least a strict inequality, $Z_{\boldsymbol{\lambda}}(B) \geq 0$, for all $B \in \mathcal{U} \setminus \{\{i\} : i \in \Omega\}$, and $\pi_{\boldsymbol{\lambda}} \leq 0$.

Proof. Since every belief function is completely characterized by its Möbius inverse, statement *(i)* is equivalent to the existence of a non-negative function $\widehat{\mu} : \mathcal{F} \to \mathbb{R}$ such that

$$\widehat{\mu}(\emptyset) = 0, \quad \sum_{A \in \mathcal{F}} \widehat{\mu}(A) = 1, \quad \text{and} \quad \widehat{Bel}(A) = \sum_{B \subseteq A} \widehat{\mu}(B), \quad \text{for every } A \in \mathcal{F},$$

further satisfying $\widehat{\mu}(\{i\}) > 0$, for all $i \in \Omega$, and

$$\mathbb{C}_{\widehat{Bel}}(\tilde{S}_1^k) = \sum_{B \in \mathcal{U}} [\tilde{S}_1^k]^{\mathbf{L}}(B)\widehat{\mu}(B) = \underline{\pi}'(S_1^k), \quad \text{for } k = 1, \dots, m.$$

Fix an enumeration of $\mathcal{U} = \{B_1, \dots, B_{2^n-1}\}$ such that $B_i = \{i\}$, for $i = 1, \dots, n$, and consider the matrices $\mathbf{A} \in \mathbb{R}^{(2(m+1)+2^n-(n+1)) \times (2^n-1)}$ and $\mathbf{B} \in \mathbb{R}^{n \times (2^n-1)}$ defined as

$$\mathbf{A} = \left(\begin{array}{c} \mathbf{C} \\ \hline \mathbf{O}_1 | -\mathbf{I}_{(2^n-(n+1))} \end{array} \right) \quad \text{and} \quad \mathbf{B} = \left(-\mathbf{I}_n \,|\mathbf{O}_2\right),$$

where $\mathbf{C} \in \mathbb{R}^{2(m+1) \times (2^n-1)}$ is defined as

$$\mathbf{C} = \begin{pmatrix} \mathbf{1}_\Omega^{\mathbf{L}}(B_1) & \cdots & \mathbf{1}_\Omega^{\mathbf{L}}(B_{2^n-1}) \\ -\mathbf{1}_\Omega^{\mathbf{L}}(B_1) & \cdots & -\mathbf{1}_\Omega^{\mathbf{L}}(B_{2^n-1}) \\ [\tilde{S}_1^1]^{\mathbf{L}}(B_1) & \cdots & [\tilde{S}_1^1]^{\mathbf{L}}(B_{2^n-1}) \\ -[\tilde{S}_1^1]^{\mathbf{L}}(B_1) & \cdots & -[\tilde{S}_1^1]^{\mathbf{L}}(B_{2^n-1}) \\ \vdots & & \vdots \\ [\tilde{S}_1^m]^{\mathbf{L}}(B_1) & \cdots & [\tilde{S}_1^m]^{\mathbf{L}}(B_{2^n-1}) \\ -[\tilde{S}_1^m]^{\mathbf{L}}(B_1) & \cdots & -[\tilde{S}_1^m]^{\mathbf{L}}(B_{2^n-1}) \end{pmatrix},$$

in which $\mathbf{I}_{(2^n-(n+1))} \in \mathbb{R}^{(2^n-(n+1)) \times (2^n-(n+1))}$ and $\mathbf{I}_n \in \mathbb{R}^{n \times n}$ are identity matrices, and $\mathbf{O}_1 \in \mathbb{R}^{(2^n-(n+1)) \times n}$ and $\mathbf{O}_2 \in \mathbb{R}^{n \times (2^n-(n+1))}$ are null matrices. Take the vector

$$\mathbf{b} = (1, -1, \underline{\pi}'(S_1^1), -\underline{\pi}'(S_1^1), \dots, \underline{\pi}'(S_1^m), -\underline{\pi}'(S_1^m), 0, \dots, 0)^T$$

with $\mathbf{b} \in \mathbb{R}^{(2(m+1)+2^n-(n+1))}$ and consider the unknown vector

$$\mathbf{x} = (\widehat{\mu}(B_1), \dots, \widehat{\mu}(B_{2^n-1}))^T$$

with $\mathbf{x} \in \mathbb{R}^{2^n-1}$. Condition *(i)* turns out to be equivalent to the solvability of the following system

$$\begin{cases} \mathbf{Ax} \leq \mathbf{b}, \\ \mathbf{Bx} < \mathbf{0}. \end{cases}$$

By the Motzkin's theorem of the alternative (Theorem 1.1) the above system is solvable if and only if for every $\mathbf{y} = (y_0, y_0', y_1, y_1', \dots, y_m, y_m', \alpha_{n+1}, \dots, \alpha_{2^n-1})^T \in \mathbb{R}^{(2(m+1)+2^n-(n+1))}$ and $\mathbf{z} = (z_1, \dots, z_n)^T \in \mathbb{R}^n$ with $\mathbf{y} \geq \mathbf{0}$ and $\mathbf{z} \geq \mathbf{0}$, none of the following conditions holds:

- $\mathbf{A}^T\mathbf{y} + \mathbf{B}^T\mathbf{z} = \mathbf{0}$, $\mathbf{z} = \mathbf{0}$ and $\mathbf{b}^T\mathbf{y} < 0$;
- $\mathbf{A}^T\mathbf{y} + \mathbf{B}^T\mathbf{z} = \mathbf{0}$, $\mathbf{z} \neq \mathbf{0}$ and $\mathbf{b}^T\mathbf{y} \leq 0$.

In turn, setting $\lambda_k = y_k - y'_k$, for $k = 0, \dots, m$, and considering $\tilde{\mathbf{y}} \in \mathbb{R}^{((m+1)+2^n-(n+1))}$, $\tilde{\mathbf{A}} \in \mathbb{R}^{((m+1)+2^n-(n+1))\times(2^n-1)}$ and $\tilde{\mathbf{b}} \in \mathbb{R}^{((m+1)+2^n-(n+1))}$, with

$$\tilde{\mathbf{y}} = (\lambda_0, \dots, \lambda_m, \alpha_{n+1}, \dots, \alpha_{2^n-1})^T \quad \text{such that } \alpha_{n+1}, \dots, \alpha_{2^n-1} \geq 0,$$

$$\tilde{\mathbf{A}} = \left(\begin{array}{c} \tilde{\mathbf{C}} \\ \hline \mathbf{O}_1 | - \mathbf{I}_{(2^n-(n+1))} \end{array}\right) \quad \text{and} \quad \tilde{\mathbf{b}} = (1, \underline{\pi}'(S_1^1), \dots, \underline{\pi}'(S_1^m), 0, \dots, 0)^T,$$

where $\tilde{\mathbf{C}} \in \mathbb{R}^{(m+1)\times(2^n-1)}$ is defined as

$$\tilde{\mathbf{C}} = \begin{pmatrix} \mathbf{1}_\Omega^{\mathbf{L}}(B_1) & \cdots & \mathbf{1}_\Omega^{\mathbf{L}}(B_{2^n-1}) \\ [\tilde{S}_1^1]^{\mathbf{L}}(B_1) & \cdots & [\tilde{S}_1^1]^{\mathbf{L}}(B_{2^n-1}) \\ \vdots & & \vdots \\ [\tilde{S}_1^m]^{\mathbf{L}}(B_1) & \cdots & [\tilde{S}_1^m]^{\mathbf{L}}(B_{2^n-1}) \end{pmatrix},$$

the above conditions can be rewritten as:

- $\tilde{\mathbf{A}}^T\tilde{\mathbf{y}} + \mathbf{B}^T\mathbf{z} = \mathbf{0}$, $\mathbf{z} = \mathbf{0}$ and $\tilde{\mathbf{b}}^T\tilde{\mathbf{y}} < 0$;
- $\tilde{\mathbf{A}}^T\tilde{\mathbf{y}} + \mathbf{B}^T\mathbf{z} = \mathbf{0}$, $\mathbf{z} \neq \mathbf{0}$ and $\tilde{\mathbf{b}}^T\tilde{\mathbf{y}} \leq 0$.

Denoting $\boldsymbol{\lambda} = (\lambda^0, \dots, \lambda^m)^T \in \mathbb{R}^{m+1}$, we have that

$$(\tilde{\mathbf{A}}^T\tilde{\mathbf{y}} + \mathbf{B}^T\mathbf{z})_i = \begin{cases} Z_{\boldsymbol{\lambda}}(B_i) - z_i, & \text{for } i = 1, \dots, n, \\ Z_{\boldsymbol{\lambda}}(B_i) - \alpha_i, & \text{for } i = n+1, \dots, 2^n - 1, \end{cases}$$

and further $\tilde{\mathbf{b}}^T\tilde{\mathbf{y}} = \pi_{\boldsymbol{\lambda}}$.

Hence, for every $\boldsymbol{\lambda} = (\lambda_0, \dots, \lambda_m)^T \in \mathbb{R}^{m+1}$, the above conditions can be rewritten as

(a') $Z_{\boldsymbol{\lambda}}(\{i\}) = 0$, for $i = 1, \dots, n$, $Z_{\boldsymbol{\lambda}}(B) \geq 0$, for all $B \in \mathcal{U} \setminus \{\{i\} : i \in \Omega\}$ and $\pi_{\boldsymbol{\lambda}} < 0$;

(b') $Z_{\boldsymbol{\lambda}}(\{i\}) \geq 0$, for $i = 1, \dots, n$, with at least a strict inequality, $Z_{\boldsymbol{\lambda}}(B) \geq 0$, for all $B \in \mathcal{U} \setminus \{\{i\} : i \in \Omega\}$, and $\pi_{\boldsymbol{\lambda}} \leq 0$.

□

To avoid heavy notation, we will sometimes omit the term "one-step" referred to the generalized arbitrage and generalized Dutch book, since we are always working in one-period setting, hence it is implied.

Recall that we interpret $\pi_{\boldsymbol{\lambda}}$ as the hypothetical price of the portfolio $\boldsymbol{\lambda}$ as if we were in a situation of completely resolving uncertainty. In this light, conditions *(ii.a)* and *(ii.b)* of previous theorem can be interpreted as two generalized forms of (one-step) arbitrage, working under partially resolving uncertainty. Avoiding condition *(ii.a)* assures that we cannot find a portfolio $\boldsymbol{\lambda}$ whose hypothetical price $\pi_{\boldsymbol{\lambda}}$ is negative (that is we are paid for it), resulting in a uniformly non-negative payoff

$Z_{\boldsymbol{\lambda}}$ in all the possible events in $\mathcal{U}$, with null value on the singletons (i.e., on those events where we have completely resolving uncertainty). Avoiding condition *(ii.b)* assures that we cannot find a portfolio $\boldsymbol{\lambda}$ whose hypothetical price $\pi_{\boldsymbol{\lambda}}$ is negative or null (that is we are paid or we do not pay anything for it), resulting in a uniformly non-negative payoff $Z_{\boldsymbol{\lambda}}$ in all the possible events in $\mathcal{U}$, with at least a strictly positive value on the singletons (i.e., on those events where we have completely resolving uncertainty).

It is immediate to see that the generalized (one-step) no-arbitrage principle expressed by statement *(ii)* of Theorem 4.5 implies the generalized (one-step) avoiding Dutch book condition in statement *(ii)* of Theorem 4.4. In particular, if a portfolio $\boldsymbol{\lambda}$ satisfies condition *(ii.a)* of Theorem 4.5 then the corresponding gain $G_{\boldsymbol{\lambda}}$ is such that $\min\limits_{B\in\mathcal{U}} G_{\boldsymbol{\lambda}}(B) > 0$, thus violating the generalized avoiding Dutch book condition.

Note 12: *The generalized (one-step) no-arbitrage principle of Theorem 4.5 is actually weaker than the classical no-arbitrage principle in one-step setting (Definition 2.4). This is due to the fact that if a portfolio $\boldsymbol{\lambda}$ gives rise to a generalized arbitrage of the form (ii.a) or (ii.b) then it also gives rise to a classical arbitrage, while a portfolio $\boldsymbol{\lambda}$ giving rise to a classical arbitrage does not generally give rise to a generalized arbitrage, as we show in the following Example 4.6.*

Let us stress that the price functional determined by the discounted Choquet expectation with respect to a $\widehat{Bel}$ with Möbius inverse $\widehat{\mu}$ as in Theorem 4.5 is generally not linear. In particular, we have that

$$\begin{aligned}
\sum_{B\in\mathcal{U}} Z_{\boldsymbol{\lambda}}(B)\widehat{\mu}(B) &= \sum_{B\in\mathcal{U}}\left(\lambda^0+\sum_{k=1}^m \lambda^k[\tilde{S}_1^k]^{\mathbf{L}}(B)\right)\widehat{\mu}(B) \\
&= \lambda^0+\sum_{k=1}^m \lambda^k\left(\sum_{B\in\mathcal{U}}[\tilde{S}_1^k]^{\mathbf{L}}(B)\widehat{\mu}(B)\right) \\
&= \lambda^0+\sum_{k=1}^m \lambda^k\mathbb{C}_{\widehat{Bel}}(\tilde{S}_1^k) = \lambda^0+\sum_{k=1}^m \lambda^k\underline{\pi}'(S_1^k) = \pi_{\boldsymbol{\lambda}}.
\end{aligned} \tag{4.13}$$

Nevertheless, considering the random variable $\lambda^0+\sum_{k=1}^m \lambda^k\tilde{S}_1^k \in \mathbb{R}^\Omega$, though $\mathbb{C}_{\widehat{Bel}}\left(\lambda^0+\sum_{k=1}^m \lambda^k\tilde{S}_1^k\right) = \lambda^0+\mathbb{C}_{\widehat{Bel}}\left(\sum_{k=1}^m \lambda^k\tilde{S}_1^k\right)$, in general we have that

$$\mathbb{C}_{\widehat{Bel}}\left(\sum_{k=1}^m \lambda^k\tilde{S}_1^k\right) \neq \sum_{k=1}^m \mathbb{C}_{\widehat{Bel}}(\lambda^k\tilde{S}_1^k) \quad \text{and} \quad \sum_{k=1}^m \mathbb{C}_{\widehat{Bel}}(\lambda^k\tilde{S}_1^k) \neq \sum_{k=1}^m \lambda^k\mathbb{C}_{\widehat{Bel}}(\tilde{S}_1^k).$$

Clearly, in the above formulas we have equalities in case $\widehat{Bel}$ reduces to a probability measure. On the other hand, in the particular case $\tilde{S}_1^h, \tilde{S}_1^k$ are comonotone and $\lambda^1, \lambda^2 \geq 0$, it holds that (see *(Ch.6)*, p.19)

$$\mathbb{C}_{\widehat{Bel}}\left(\lambda^1\tilde{S}_1^h+\lambda^2\tilde{S}_1^k\right) = \lambda^1\mathbb{C}_{\widehat{Bel}}(\tilde{S}_1^h)+\lambda^2\mathbb{C}_{\widehat{Bel}}(\tilde{S}_1^k) = \lambda^1\underline{\pi}'(S_1^h)+\lambda^2\underline{\pi}'(S_1^k).$$

Furthermore, denoting by $\widehat{Pl}$ the dual plausibility function of $\widehat{Bel}$, we have that for a generic random variable $X \in \mathbb{R}^\Omega$, it holds that

$$(1+r)^{-1}\mathbb{C}_{\widehat{Bel}}(X) \leq (1+r)^{-1}\mathbb{C}_{\widehat{Pl}}(X), \tag{4.14}$$

i.e., the two values above should be interpreted as lower and upper prices.

The following example shows a lower price assessment violating the generalized (one-step) no-arbitrage principle expressed in Theorem 4.5.

Example 4.5 *Let $\Omega = \{1,2,3,4\}$, $\mathcal{F} = \mathcal{P}(\Omega)$ and consider three contracts whose payoffs in euros at time $t = 1$ are*

Ω	1	2	3	4
S_1^1	10	10	20	20
S_1^2	0	10	0	10
S_1^3	10	30	20	40

Assume that the lower prices at time $t = 0$ are fixed to $\underline{\pi}'(S_1^1) = 15$, $\underline{\pi}'(S_1^2) = 5$ and $\underline{\pi}'(S_1^3) = 20$ and that the risk-free interest rate is $r = 0$, so we have $\tilde{S}_1^k = S_1^k$, for $k = 1,2,3$.

This lower price assessment violates the generalized no-arbitrage principle of Theorem 4.5 as, in particular, it violates the generalized avoiding Dutch book condition expressed in Theorem 4.4. Indeed, every belief function $\widehat{Bel}$ on $\mathcal{F}$ induces a Choquet expectation functional on $\mathbb{R}^\Omega$ which is positively homogeneous and superadditive, therefore, assuming $\mathbb{C}_{\widehat{Bel}}(S_1^k) = \underline{\pi}'(S_1^k)$, for $k = 1,2,3$, it should be, as $S_1^3 = S_1^1 + 2S_1^2$,

$$\mathbb{C}_{\widehat{Bel}}(S_1^3) = \mathbb{C}_{\widehat{Bel}}(S_1^1 + 2S_1^2) \geq \mathbb{C}_{\widehat{Bel}}(S_1^1) + 2\mathbb{C}_{\widehat{Bel}}(S_1^2) = 25.$$

Denoting $\mathcal{U} = \mathcal{F} \setminus \{\emptyset\}$ and omitting braces and commas to have a lighter set notation, if we consider the portfolio $\boldsymbol{\lambda} = (0,-1,-2,1)^T$ we have that $\pi_{\boldsymbol{\lambda}} = -5$ and

$\mathcal{U}$	1	2	3	4	12	13	14	23	24	34	123	124	134	234	1234
$[S_1^1]^{\mathbf{L}}$	10	10	20	20	10	10	10	10	10	20	10	10	10	10	10
$[S_1^2]^{\mathbf{L}}$	0	10	0	10	0	0	0	0	10	0	0	0	0	0	0
$[S_1^3]^{\mathbf{L}}$	10	30	20	40	10	10	10	20	30	20	10	10	10	20	10
$Z_{\boldsymbol{\lambda}}$	0	0	0	0	0	0	0	10	0	0	0	0	0	10	0
$G_{\boldsymbol{\lambda}}$	5	5	5	5	5	5	5	15	5	5	5	5	5	15	5

Hence, since $\min_{B \in \mathcal{U}} G_{\boldsymbol{\lambda}}(B) > 0$, the generalized avoiding Dutch book condition is not satisfied, therefore there is no belief function $\widehat{Bel}$ such that $\mathbb{C}_{\widehat{Bel}}$ agrees with the assessed lower prices. Moreover, the same $\boldsymbol{\lambda}$ shows that we have a generalized arbitrage in the form of (ii.a) of Theorem 4.5, since $Z_{\boldsymbol{\lambda}}(\{i\}) = 0$, for $i = 1,\dots,4$, $Z_{\boldsymbol{\lambda}}(B) \geq 0$, for all $B \in \mathcal{U} \setminus \{\{i\} : i \in \Omega\}$ and $\pi_{\boldsymbol{\lambda}} < 0$.

On the other hand, taking the portfolio $\boldsymbol{\lambda}' = (0,-2,-10,4)^T$ we have that $\pi_{\boldsymbol{\lambda}'} = 0$ and

$\mathcal{U}$	1	2	3	4	12	13	14	23	24	34	123	124	134	234	1234
$Z_{\boldsymbol{\lambda}'}$	20	0	40	20	20	20	20	60	0	40	20	20	20	60	20

therefore, we have a generalized arbitrage in the form of (ii.b) of Theorem 4.5, since $Z_{\boldsymbol{\lambda}'}(\{i\}) \geq 0$, for $i = 1,\dots,4$, with at least a strict inequality, $Z_{\boldsymbol{\lambda}'}(B) \geq 0$, for all $B \in \mathcal{U} \setminus \{\{i\} : i \in \Omega\}$ and $\pi_{\boldsymbol{\lambda}'} \leq 0$.

♦

The following example shows a lower price assessment violating the classical (one-step) no-arbitrage principle but not the (one-step) generalized no-arbitrage principle.

Example 4.6 *Let Ω, $\mathcal{F}$, r, and S_1^1, S_1^2, S_1^3 as in Example 4.5. Consider the lower price assessment $\underline{\pi}'(S_1^1) = 15$, $\underline{\pi}'(S_1^2) = 5$ and $\underline{\pi}'(S_1^3) = 26$. Such an assessment violates the classical (one-step) no-arbitrage principle, indeed, every probability measure Q on $\mathcal{F}$ gives rise to a positive, linear and normalized functional $\mathbb{E}_Q$ on $\mathbb{R}^\Omega$. Hence, assuming $\mathbb{E}_Q(S_1^k) = \underline{\pi}'(S_1^k)$, for $k = 1,2,3$, it should be, as $S_1^3 = S_1^1 + 2S_1^2$,*

$$\mathbb{E}_Q(S_1^3) = \mathbb{E}_Q(S_1^1 + 2S_1^2) = \mathbb{E}_Q(S_1^1) + 2\mathbb{E}_Q(S_1^2) = 25.$$

On the other hand, there exists a belief function $\widehat{Bel}$ on $\mathcal{F}$ which is strictly positive on $\mathcal{F} \setminus \{\emptyset\}$, whose corresponding Choquet expectation functional $\mathbb{C}_{\widehat{Bel}}$ agrees with the given lower price assessment. For instance, we can take the $\widehat{Bel}$ whose Möbius inverse $\widehat{\mu}$ is such that

$\mathcal{U}$	1	2	3	4	12	13	14	23	24	34	123	124	134	234	1234
$[S_1^1]^{\mathbf{L}}$	10	10	20	20	10	10	10	10	10	20	10	10	10	10	10
$[S_1^2]^{\mathbf{L}}$	0	10	0	10	0	0	0	0	10	0	0	0	0	0	0
$[S_1^3]^{\mathbf{L}}$	10	30	20	40	10	10	10	20	30	20	10	10	10	20	10
$\widehat{\mu}$	$\frac{2}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{4}{10}$	$\frac{1}{10}$	0	0	$\frac{1}{10}$	0	0	0	0	0	0	0

For such a $\widehat{Bel}$ we have that

$$\begin{aligned}
\mathbb{C}_{\widehat{Bel}}(S_1^1) &= 10 \cdot \frac{2}{10} + (10+20+10+10) \cdot \frac{1}{10} + 20 \cdot \frac{4}{10} = 15, \\
\mathbb{C}_{\widehat{Bel}}(S_1^2) &= 10 \cdot \frac{1}{10} + 10 \cdot \frac{4}{10} = 5, \\
\mathbb{C}_{\widehat{Bel}}(S_1^3) &= 10 \cdot \frac{2}{10} + (30+20+10+20) \cdot \frac{1}{10} + 40 \cdot \frac{4}{10} = 26.
\end{aligned}$$

Hence, by Theorem 4.5 we cannot find a portfolio $\boldsymbol{\lambda}$ giving rise to a generalized arbitrage in the form of (ii.a) or (ii.b). On the other hand, the portfolio $\boldsymbol{\lambda} = (0,1,2,-1)^T$ gives rise to a classical arbitrage since $\pi_{\boldsymbol{\lambda}} = -1 < 0$ and

Ω	1	2	3	4
S_1^1	10	10	20	20
S_1^2	0	10	0	10
S_1^3	10	30	20	40
$\lambda^0 + \sum_{k=1}^3 \lambda^k S_1^k$	0	0	0	0

♦

Remark 4. A non-linear pricing rule defined as a discounted Choquet expectation with respect to a convex (concave) capacity has been characterized in literature as we summarized in Section 3.1.2. In particular, Chateauneuf et al. (1996); Cerreia-Vioglio et al. (2015); Chateauneuf and Cornet (2022b) proposed a Choquet pricing rule satisfying a form of put-call parity. Their functional can be interpreted as an upper pricing rule and those proposed by Cerreia-Vioglio et al. (2015) and Chateauneuf and

Cornet (2022b) allow to model frictions in the market. Our model results to be a more specific case of the quoted works since they propose an axiomatic characterization of the non-linear pricing rule without giving a normative justification in terms of no-arbitrage condition of the proposed pricing rule. On the contrary, our generalized one-step no-arbitrage condition has a clear interpretation like the classical one (see Section 2.1) and it is equivalent to the existence of a (possibly not unique) *completely monotone* discounted Choquet expectation that can be interpreted as a lower pricing rule still allowing for frictions in the market.

The completeness of the market is, generally, not achieved, since Theorem 4.5 does not guarantee the uniqueness of $\widehat{Bel}$. However, we stress that, if $\{\mathbf{1}_B : B \in \mathcal{U}\} \subseteq \mathcal{G}$ and $\underline{\pi}' : \mathcal{G} \to \mathbb{R}$ satisfies the generalized no-arbitrage principle, then there exists a *unique* $\widehat{Bel}$, positive on $\mathcal{F} \setminus \{\emptyset\}$, such that the corresponding discounted Choquet expectation functional on $\mathbb{R}^\Omega$ agrees with $\underline{\pi}'$. The payoffs in $\{\mathbf{1}_B : B \in \mathcal{U}\}$ can be considered as *generalized Arrow-Debreu securities* (see, e.g., Dybvig and Ross, 1989; Pliska, 1997; Cerný, 2009), working under partially resolving uncertainty. Indeed, the family of lower prices $\{\underline{\pi}'(\mathbf{1}_B) : B \in \mathcal{U}\}$ uniquely determines $\widehat{Bel}$ since, for every $B \in \mathcal{U}$, we have that

$$\underline{\pi}'(\mathbf{1}_B) = (1+r)^{-1}\widehat{Bel}(B), \tag{4.15}$$

therefore such lower prices can be dubbed *generalized Arrow-Debreu lower prices.*

Note 13: *The wider framework of* 2*-monotone capacities could be considered instead of Dempster-Shafer theory. The reason why we stick to Dempster-Shafer theory is that the generalized one-step no-arbitrage condition we get has a clear interpretation and its connection to the classical setting is evident. Like the classical no-arbitrage principle, our generalized no-arbitrage principle has a normative purpose: respecting it we derive a (non-necessarily unique) lower pricing rule that allows us to price securities taking care of bid-ask spreads. Besides equivalence,* $\widehat{Bel}$ *has no other relation with Bel nor with market agents' utility functions as it is only determined by the generalized no-arbitrage principle. Switching to the* 2*-monotone setting, though mathematically possible, makes the no-arbitrage condition much more involved and its interpretation is difficult to justify from a normative point of view.*

4.4 Equivalent inner approximating Choquet martingale belief functions

We turn back to the lower envelope $\underline{Q}$ of the class $\mathcal{Q}$ of equivalent martingale measures induced by the n-nomial market model characterized in Section 4.1. Recall that in this context we have only one risky asset whose price process is $\{S_0^1, S_1^1\}$. In what follows we assume that the lower price of the risky asset is $\underline{\pi}'(S_1^1) = S_0^1$, which is justified by the fact that $\min\limits_{Q \in \mathrm{cl}(\mathcal{Q})} (1+r)^{-1}\mathbb{E}_Q(S_1^1) = S_0^1$. Moreover, here the "real-world" probability P can play the role of the "real-world" belief function introduced in Section 4.3.

As already pointed out in Corollary 4.1, $\underline{Q}$ is a belief function that we know is not positive over $\mathcal{U}$, for $n > 2$. Moreover, we have that assessing $\underline{\pi}'(S_1^1) = S_0^1$ is generally not consistent with the lower pricing rule obtained as the discounted

Choquet expectation with respect to $\underline{Q}$. Indeed, if we define $\underline{\pi}' : \mathcal{G} \to \mathbb{R}$ on $\mathcal{G} = \{S_1^1\} \cup \{\mathbf{1}_B : B \in \mathcal{U}\}$ such that

$$\underline{\pi}'(S_1^1) = S_0^1 \quad \text{and} \quad \underline{\pi}'(\mathbf{1}_B) = (1+r)^{-1}\underline{Q}(B), \quad \text{for all } B \in \mathcal{U}, \tag{4.16}$$

we get that $\underline{\pi}'$ violates both the generalized avoiding Dutch book condition and the generalized no-arbitrage principle, as Example 4.7 shows.

Example 4.7 *Consider $\Omega = \{1,2,3,4\}$, $m_1 = 4$, $m_2 = 2$, $m_3 = \frac{1}{2}$, $m_4 = \frac{1}{4}$, $1+r = 1$, $\mathcal{Q}$ and $\underline{Q}$ of Example 4.1. Let $S_0^1 = 100$ and take the lower price assessment $\underline{\pi}'$ on $\mathcal{G} = \{S_1^1\} \cup \{\mathbf{1}_B : B \in \mathcal{U}\}$ defined as in (4.16).*

Denote by $\lambda^{S^0}, \lambda^{S^1}$ the numbers of units of the risk-free and risky assets, and by $\lambda^i, \lambda^{ij}, \lambda^{ijk}, \lambda^{1234}$ those associated with $\mathbf{1}_{\{i\}}, \mathbf{1}_{\{i,j\}}, \mathbf{1}_{\{i,j,k\}}, \mathbf{1}_\Omega$, respectively, in a portfolio $\boldsymbol{\lambda}$ on $\mathcal{G}$. If we take

λ^{S^0}	λ^{S^1}	λ^1	λ^2	λ^3	λ^4	λ^{12}	λ^{13}	λ^{14}	λ^{23}	λ^{24}	λ^{34}	λ^{123}	λ^{124}	λ^{134}	λ^{234}	λ^{1234}
0	$-\frac{4}{25}$	7	10	-10	-10	8	10	5	-1	-5	-10	5	10	9	-5	10

we get that $\pi_{\boldsymbol{\lambda}} = -7$ and further

$\mathcal{U}$	1	2	3	4	12	13	14	23	24	34	123	124	134	234	1234
$Z_{\boldsymbol{\lambda}}$	0	0	0	0	1	26	30	1	6	0	7	16	15	1	6

In turn, since $\pi_{\boldsymbol{\lambda}} < 0$ and $G_{\boldsymbol{\lambda}}$ is defined by (4.11), it follows that $\min_{B \in \mathcal{U}} G_{\boldsymbol{\lambda}}(B) > 0$, thus $\boldsymbol{\lambda}$ violates the avoiding Dutch book condition. We also notice that this portfolio gives rise to a generalized arbitrage of type (ii.a) in Theorem 4.5, since $Z_{\boldsymbol{\lambda}}(\{i\}) = 0$, for $i = 1, \ldots, 4$, $Z_{\boldsymbol{\lambda}}(B) \geq 0$, for all $B \in \mathcal{U} \setminus \{\{i\} : i \in \Omega\}$ and $\pi_{\boldsymbol{\lambda}} < 0$.

We stress that $\underline{\pi}'$ fails the generalized avoiding Dutch book condition (and, therefore, the generalized no-arbitrage principle) since S_1^1 is mispriced as it should be $\underline{\pi}'(S_1^1) = \mathbb{C}_{\underline{Q}}(S_1^1) = \frac{5400}{105}$ to be consistent with the generalized Arrow-Debreu lower prices.

Furthermore, if we take the portfolio $\tilde{\boldsymbol{\lambda}}$ on $\mathcal{G}$ with entries

$\tilde{\lambda}^{S^0}$	$\tilde{\lambda}^{S^1}$	$\tilde{\lambda}^1$	$\tilde{\lambda}^2$	$\tilde{\lambda}^3$	$\tilde{\lambda}^4$	$\tilde{\lambda}^{12}$	$\tilde{\lambda}^{13}$	$\tilde{\lambda}^{14}$	$\tilde{\lambda}^{23}$	$\tilde{\lambda}^{24}$	$\tilde{\lambda}^{34}$	$\tilde{\lambda}^{123}$	$\tilde{\lambda}^{124}$	$\tilde{\lambda}^{134}$	$\tilde{\lambda}^{234}$	$\tilde{\lambda}^{1234}$
0	$-\frac{1}{5}$	10	10	10	10	10	10	10	10	10	6	10	10	10	-5	10

we get that $\pi_{\tilde{\boldsymbol{\lambda}}} = 0$ and

$\mathcal{U}$	1	2	3	4	12	13	14	23	24	34	123	124	134	234	1234
$Z_{\tilde{\boldsymbol{\lambda}}}$	0	25	51	56	0	30	35	15	20	16	10	15	15	0	5

It is easily seen that the portfolio $\tilde{\boldsymbol{\lambda}}$ gives rise to a generalized arbitrage of type (ii.b) in Theorem 4.5, since $Z_{\tilde{\boldsymbol{\lambda}}}(\{i\}) \geq 0$, for $i = 1, \ldots, 4$, with at least a strict inequality, $Z_{\tilde{\boldsymbol{\lambda}}}(B) \geq 0$, for all $B \in \mathcal{U} \setminus \{\{i\} : i \in \Omega\}$, and $\pi_{\tilde{\boldsymbol{\lambda}}} \leq 0$.

♦

Let us stress that another problem related to $\underline{Q}$ is its non-positivity on $\mathcal{U}$. To see this, let $\mathcal{G}' = \{\mathbf{1}_B : B \in \mathcal{U}\}$ and consider the restriction of the lower price assessment (4.16) defined as $\underline{\pi}'' = \underline{\pi}'_{|\mathcal{G}'}$. In this case, since the lower price assessment is consistent with the discounted Choquet expectation with respect to $\underline{Q}$, no generalized arbitrage in the form of *(ii.a)* in Theorem 4.5 can be built, as it would imply the violation of the generalized avoiding Dutch book condition. On the other hand, a generalized arbitrage in the form of *(ii.b)* in Theorem 4.5 can be found as the following Example 4.8 shows.

Example 4.8 *Let Ω, m_1, m_2, m_3, m_4, $1+r$, $\mathcal{Q}$, $\underline{Q}$, and S_0^1 be defined as in Example 4.7. Denote by $\hat{\boldsymbol{\lambda}}$ a portfolio on $\mathcal{G}'$, whose components are $\hat{\lambda}^{S^0}$, and $\hat{\lambda}^i, \hat{\lambda}^{ij}, \hat{\lambda}^{ijk}, \hat{\lambda}^{1234}$, referring to $\mathbf{1}_{\{i\}}, \mathbf{1}_{\{i,j\}}, \mathbf{1}_{\{i,j,k\}}, \mathbf{1}_{\Omega}$, respectively. If we take*

$\hat{\lambda}^{S^0}$	$\hat{\lambda}^1$	$\hat{\lambda}^2$	$\hat{\lambda}^3$	$\hat{\lambda}^4$	$\hat{\lambda}^{12}$	$\hat{\lambda}^{13}$	$\hat{\lambda}^{14}$	$\hat{\lambda}^{23}$	$\hat{\lambda}^{24}$	$\hat{\lambda}^{34}$	$\hat{\lambda}^{123}$	$\hat{\lambda}^{124}$	$\hat{\lambda}^{134}$	$\hat{\lambda}^{234}$	$\hat{\lambda}^{1234}$
0	10	10	10	10	-10	10	10	10	10	-10	-10	10	10	-10	10

we get that $\pi_{\hat{\boldsymbol{\lambda}}} = 0$ and also

$\mathcal{U}$	1	2	3	4	12	13	14	23	24	34	123	124	134	234	1234
$Z_{\hat{\boldsymbol{\lambda}}}$	40	20	20	40	0	20	40	0	20	0	0	20	20	0	10

One immediately observes that the portfolio $\hat{\boldsymbol{\lambda}}$ gives rise to a generalized arbitrage of type (ii.b) in Theorem 4.5.

♦

The previous examples show that if we seek a positive belief function whose discounted Choquet expectation is consistent with the generalized no-arbitrage principle and the assessment $\underline{\pi}'(S_1^1) = S_0^1$, then we must depart from $\underline{Q}$.

It is easily seen that every equivalent martingale measure $Q \in \mathcal{Q}$ is a belief function, positive on $\mathcal{U}$, nevertheless, the choice of a particular Q_0 in the class $\mathcal{Q}$ is a problematic task, as one needs to provide a reasonable choice criterion. For instance, a possible choice is $Q_0 = \frac{1}{|\mathrm{ext}(\mathrm{cl}(\mathcal{Q}))|}\sum_{Q\in\mathrm{ext}(\mathrm{cl}(\mathcal{Q}))} Q$, which belongs to $\mathcal{Q}$ since it is a strict convex combination of elements of $\mathrm{ext}(\mathrm{cl}(\mathcal{Q}))$.

Once an equivalent martingale measure Q_0 has been chosen, some of the information contained in the class $\mathcal{Q}$ can be preserved if we consider (see, e.g., Huber, 1981; Walley, 1991) the *ϵ-contamination* of Q_0 with respect to $\mathrm{cl}(\mathcal{Q})$, where $\epsilon \in (0,1)$. This amounts to consider the closed subset of $\mathcal{Q}$ given by

$$\mathcal{Q}_\epsilon = \{Q \in \mathcal{Q} : Q = (1-\epsilon)Q_0 + \epsilon Q', Q' \in \mathrm{cl}(\mathcal{Q})\}, \tag{4.17}$$

whose lower envelope $\underline{Q}_\epsilon = \min \mathcal{Q}_\epsilon$ is defined, for every $A \in \mathcal{F}$, as

$$\underline{Q}_\epsilon(A) = (1-\epsilon)Q_0(A) + \epsilon\underline{Q}(A). \tag{4.18}$$

In particular, since $\underline{Q}_\epsilon$ is the strict convex combination of the two belief functions Q_0 and $\underline{Q}$, we have that $\underline{Q}_\epsilon$ is a belief function which, in turn, is strictly positive over $\mathcal{U}$, as Q_0 is.

The idea is to directly use $\underline{Q}_\epsilon$ in order to derive a lower pricing rule through a discounted Choquet expectation. In the light of previous section, the obtained lower pricing rule would be acceptable if it satisfied the generalized (one-step) no-arbitrage condition expressed by Theorem 4.5.

The following Example 4.9 shows that, even if $\underline{Q}_\epsilon$ is a positive belief function, the corresponding lower pricing rule obtained as discounted Choquet expectation is not consistent with $\underline{\pi}'(S_1^1) = S_0^1$. In particular, this means that we can build portfolios resulting in generalized arbitrage opportunities and generalized Dutch books. Indeed, the lower pricing rule with respect to $\underline{Q}_\epsilon$ can be expressed as a convex combination of the discounted Choquet expectations with respect to $\underline{Q}$ and Q_0. The problem with $\underline{Q}_\epsilon$ is that $\mathrm{cl}(\mathcal{Q}_\epsilon)$ is strictly contained in $\mathrm{core}(\underline{Q}_\epsilon)$ (as discussed in *(Ch.8)* p.19).

Example 4.9 *Consider $\Omega = \{1, 2, 3, 4\}$, $m_1 = 4$, $m_2 = 2$, $m_3 = \frac{1}{2}$, $m_4 = \frac{1}{4}$, $1 + r = 1$, $\mathcal{Q}$ and $\underline{Q}$ of Example 4.1, and let Q_0 be an arbitrary element of $\mathcal{Q}$. Take $\epsilon \in (0, 1)$ and consider the ϵ-contamination class of Q_0 with respect to* $\mathrm{cl}(\mathcal{Q})$*, whose lower envelope is $\underline{Q}_\epsilon = (1 - \epsilon)Q_0 + \epsilon\underline{Q}$.*

The lower pricing rule obtained as the discounted Choquet expectation with respect $\underline{Q}_\epsilon$ is not consistent with the assessment $\underline{\pi}'(S_1^1) = S_0^1$ as it should be

$$(1 + r)^{-1}\mathbb{C}_{\underline{Q}_\epsilon}(S_1^1) = S_0^1,$$

which is equivalent, since $S_0^1 = s > 0$, to

$$\mathbb{C}_{\underline{Q}_\epsilon}\left(\frac{S_1^1}{S_0^1}\right) = 1 + r.$$

Nevertheless, due to property (Ch.2) (p.19) of the Choquet integral, we have that

$$\begin{aligned}
\mathbb{C}_{\underline{Q}_\epsilon}\left(\frac{S_1^1}{S_0^1}\right) &= \mathbb{C}_{(1-\epsilon)Q_0+\epsilon\underline{Q}}\left(\frac{S_1^1}{S_0^1}\right) = (1-\epsilon)\mathbb{C}_{Q_0}\left(\frac{S_1^1}{S_0^1}\right) + \epsilon\mathbb{C}_{\underline{Q}}\left(\frac{S_1^1}{S_0^1}\right) \\
&= (1-\epsilon) + \epsilon\frac{54}{105} < 1 + r,
\end{aligned}$$

since $\mathbb{C}_{\underline{Q}}\left(\frac{S_1^1}{S_0^1}\right) = \frac{54}{105}$.

♦

Note 14: *The result in Example 4.9 does not depend on Q_0. Indeed, for every choice of $Q_0 \in \mathcal{Q}$, then $\underline{Q}_\epsilon$ gives rise to a lower pricing rule that assigns positive lower price to every security with non-negative and non-null payoff but is not consistent with $\underline{\pi}'(S_1^1) = S_0^1$. Indeed, the issue of ϵ-contamination in this case is due to the fact that* $\mathrm{cl}(\mathcal{Q})$ *is strictly contained in* $\mathrm{core}(\underline{Q})$*, so we have a situation like that described in property (Ch.8) (p.19) for the corresponding Choquet expectation $\mathbb{C}_{\underline{Q}}$.*

A possible way to keep the assessment $\underline{\pi}'(S_1^1) = S_0^1$ and fulfill the generalized no-arbitrage principle is to ϵ-contaminate Q_0 (Huber, 1981; Moral, 2018) with an inner approximation $\widehat{Bel}$ of $\underline{Q}$, similarly to the approach of Miranda et al. (2022), requiring, in addition, to satisfy the generalized avoiding Dutch book condition. We would define, for $\epsilon \in (0, 1)$,

$$\widehat{Bel}_\epsilon = (1 - \epsilon)Q_0 + \epsilon\widehat{Bel}. \tag{4.19}$$

The belief functions $\widehat{Bel}$ and $\widehat{Bel}_\epsilon$ will be referred to as *inner approximating one-step Choquet martingale belief function* and *equivalent inner approximating one-step Choquet martingale belief function*, according to the following definition.

Definition 4.2 *A belief function $\widehat{Bel}$ on $\mathcal{F}$ is called:*

- *an* **inner approximation** *for $\underline{Q}$ if, for every $A \in \mathcal{F}$, it holds that*

$$\underline{Q}(A) \leq \widehat{Bel}(A);$$

- *a one-step* **Choquet martingale belief function** *if*

$$(1+r)^{-1}\mathbb{C}_{\widehat{Bel}}(S_1^1) = S_0^1;$$

- *an* **inner approximating one-step Choquet martingale belief function** *for $\underline{Q}$ if it is both an inner approximation for $\underline{Q}$ and a one-step Choquet martingale belief function;*

- *an* **equivalent inner approximating one-step Choquet martingale belief function** *for $\underline{Q}$ if it is an inner approximating one-step Choquet martingale belief function for $\underline{Q}$ and $\widehat{Bel} \sim P$.*

Our goal is to select a belief function $\widehat{Bel} \in \mathbf{B}(\Omega, \mathcal{F})$ which is an inner approximating one-step Choquet martingale belief function for $\underline{Q}$ and is closest to $\underline{Q}$ with respect to a suitable distance d defined on the set $\mathbf{B}(\Omega, \mathcal{F})$.

Note 15: *Our approach differs from that of Montes et al. (2018); Miranda et al. (2021); Montes et al. (2019) since they look for an outer approximation, i.e., a belief function such that $\widehat{Bel} \leq \underline{Q}$, where comparisons are pointwise on $\mathcal{F}$. On the contrary, we follow the approach of Miranda et al. (2022), where they look for a 2-monotone and belief function inner approximating a (coherent) lower probability. Our choice of an inner approximation, rather than an outer approximation, is due to the fact that the latter would imply a greater dilation in the price interval with respect to the inner approximation. Moreover, an outer approximation $\widehat{Bel}$ would induce* $\mathrm{core}(\widehat{Bel})$ *that contains* $\mathrm{core}(\underline{Q})$*, and we already know that* $\mathrm{cl}(\mathcal{Q}) \subset \mathrm{core}(\underline{Q})$*, so the martingale property for the stock cannot be enforced.*

Moreover, since the entire focus of the study is under Dempster-Shafer uncertainty (due to the fact that $\underline{Q}$ is a belief function), the aim is to arrive to a consistent lower pricing rule inside of the same framework. However, the same process could be carried out in the more general framework of 2-monotone capacities, as shown in Montes et al. (2018, 2019).

The trivial solution to the problem of inner approximating $\underline{Q}$ should be taking a $Q \in \mathrm{cl}(\mathcal{Q})$. It is easy to see that every $Q \in \mathrm{cl}(\mathcal{Q})$ is an inner approximating one-step Choquet martingale belief function for $\underline{Q}$, while every $Q \in \mathcal{Q}$ is an equivalent inner approximating one-step Choquet martingale belief function for $\underline{Q}$. Thus, a possible procedure would be to choose a probability measure $Q \in \mathcal{Q}$. Nevertheless, this procedure suffers from the well-known problem of providing a criterion to choose a $Q \in \mathcal{Q}$ and further it does not allow to model frictions in the market as bid-ask spreads. Hence, we avoid this trivial case and look for a non-additive equivalent inner approximating one-step Choquet martingale belief function.

In order to choose an approximation, following Montes et al. (2018, 2019); Miranda et al. (2022), two possible choices for the distance d are

$$d_1(Bel_1, Bel_2) = \sum_{A \in \mathcal{F}} |Bel_1(A) - Bel_2(A)|, \tag{4.20}$$

$$d_2(Bel_1, Bel_2) = \sum_{A \in \mathcal{F}} (Bel_1(A) - Bel_2(A))^2, \tag{4.21}$$

that can be seen as distances induced by the L^1 and L^2 norms on $[0,1]^{\mathcal{F}}$. In particular, d_2 is the squared Euclidean distance.

Thus, for a fixed distance d, an optimal inner approximating one-step Choquet martingale belief function $\widehat{Bel}$ for $\underline{Q}$ can be found by solving the following optimization problem:

$$\begin{gathered}\text{minimize } d(\widehat{Bel}, \underline{Q})\\ \text{subject to:}\\ \begin{cases} \widehat{Bel}(A) \geq \underline{Q}(A), & \text{for every } A \in \mathcal{F},\\ (1+r)^{-1}\mathbb{C}_{\widehat{Bel}}(S_1^1) = S_0^1, & \\ \widehat{Bel} \in \mathbf{B}(\Omega, \mathcal{F}). & \end{cases}\end{gathered} \tag{4.22}$$

The searched $\widehat{Bel}$ is completely characterized by its Möbius inverse $\widehat{\mu}$ that must satisfy $\widehat{\mu}(\emptyset) = \widehat{Bel}(\emptyset) = \underline{Q}(\emptyset) = 0$. Moreover, since $S_0^1 = s_0 > 0$, it holds that

$$(1+r)^{-1}\mathbb{C}_{\widehat{Bel}}(S_1^1) = S_0^1$$

is equivalent to

$$\mathbb{C}_{\widehat{Bel}}\left(\frac{S_1^1}{S_0^1}\right) = 1 + r,$$

where

$$\mathbb{C}_{\widehat{Bel}}\left(\frac{S_1^1}{S_0^1}\right) = \sum_{B \in \mathcal{U}} \left(\frac{S_1^1}{S_0^1}\right)^{\mathbf{L}}(B)\widehat{\mu}(B) = \sum_{i=1}^{n} m_i \left(\sum_{\{i\} \subseteq B \subseteq \{1,\ldots,i\}} \widehat{\mu}(B)\right).$$

Hence, the above problem (4.22) is equivalent to the following optimization problem with linear constraints, whose unknowns are the values of $\widehat{\mu}$ on $\mathcal{U}$:

$$\begin{gathered}\text{minimize } d(\widehat{Bel}, \underline{Q})\\ \text{subject to:}\\ \begin{cases} \sum_{\emptyset \neq B \subseteq A} \widehat{\mu}(B) \geq \underline{Q}(A), & \text{for every } A \in \mathcal{U},\\ \sum_{i=1}^{n} m_i \left(\sum_{\{i\} \subseteq B \subseteq \{1,\ldots,i\}} \widehat{\mu}(B)\right) = 1 + r, & \\ \sum_{B \in \mathcal{U}} \widehat{\mu}(B) = 1, & \\ \widehat{\mu}(B) \geq 0, & \text{for every } B \in \mathcal{U}. \end{cases}\end{gathered} \tag{4.23}$$

It is easily seen that every $Q \in \mathrm{cl}(\mathcal{Q})$ gives rise to a Möbius inverse that satisfies all the constraints in problem (4.23). Hence, the feasible region of problem (4.23) is a non-empty convex compact subset of $\mathbb{R}^{2^n-1}$, endowed with the product topology.

We notice that, as already pointed out in Montes et al. (2018, 2019), if we consider the distance d_1 and take into account that $\widehat{Bel} \geq \underline{Q}$, we have that

$$\begin{aligned} d_1(\widehat{Bel}, \underline{Q}) &= \sum_{A \in \mathcal{U}} \left[\left(\sum_{\emptyset \neq B \subseteq A} \widehat{\mu}(B) \right) - \underline{Q}(A) \right] \\ &= \sum_{A \in \mathcal{U}} 2^{|\Omega \setminus A|} \widehat{\mu}(A) - \sum_{A \in \mathcal{U}} \underline{Q}(A), \end{aligned} \tag{4.24}$$

where $\sum\limits_{A \in \mathcal{U}} \underline{Q}(A)$ is a constant, since $\underline{Q}$ is given. Therefore, problem (4.23) reduces to a linear programming problem.

The following example shows the computation of an equivalent inner approximating one-step Choquet martingale belief function, relying on the distance d_1.

Example 4.10 *Consider $\Omega = \{1,2,3,4\}$, $m_1 = 4$, $m_2 = 2$, $m_3 = \frac{1}{2}$, $m_4 = \frac{1}{4}$, $1 + r = 1$, and $\mathcal{Q}$ and $\underline{Q}$ of Example 4.1.*

An inner approximating one-step Choquet martingale belief function $\widehat{Bel}$ minimizing the d_1 distance is reported below

$\mathcal{F}$	$\emptyset$	1	2	3	4	12	13	14	23	24	34	123	124	134	234	Ω
$\underline{Q}$	0	0	0	0	0	$\frac{15}{105}$	0	0	0	0	$\frac{60}{105}$	$\frac{21}{105}$	$\frac{15}{105}$	$\frac{60}{105}$	$\frac{84}{105}$	1
μ	0	0	0	0	0	$\frac{15}{105}$	0	0	0	0	$\frac{60}{105}$	$\frac{6}{105}$	0	0	$\frac{24}{105}$	0
$\widehat{\mu}$	0	$\frac{21}{105}$	0	0	0	0	0	0	0	0	$\frac{60}{105}$	0	0	0	$\frac{24}{105}$	0
$\widehat{Bel}$	0	$\frac{21}{105}$	0	0	0	$\frac{21}{105}$	$\frac{21}{105}$	$\frac{21}{105}$	0	0	$\frac{60}{105}$	$\frac{21}{105}$	$\frac{21}{105}$	$\frac{81}{105}$	$\frac{84}{105}$	1

for which we have that $d_1(\widehat{Bel}, \underline{Q}) = \frac{96}{105}$.

Define $Q_0 = \frac{1}{|\mathrm{ext}(\mathrm{cl}(\mathcal{Q}))|} \sum_{Q \in \mathrm{ext}(\mathrm{cl}(\mathcal{Q}))} Q$, whose values are reported below

$\mathcal{F}$	$\emptyset$	1	2	3	4	12	13	14	23	24	34	123	124	134	234	Ω
$Q_{1,3}$	0	$\frac{15}{105}$	0	$\frac{90}{105}$	0	$\frac{15}{105}$	1	$\frac{15}{105}$	$\frac{90}{105}$	0	$\frac{90}{105}$	1	$\frac{15}{105}$	1	$\frac{90}{105}$	1
$Q_{1,4}$	0	$\frac{21}{105}$	0	0	$\frac{84}{105}$	$\frac{21}{105}$	$\frac{21}{105}$	1	0	$\frac{84}{105}$	$\frac{84}{105}$	$\frac{21}{105}$	1	1	$\frac{84}{105}$	1
$Q_{2,3}$	0	0	$\frac{35}{105}$	$\frac{70}{105}$	0	$\frac{35}{105}$	$\frac{70}{105}$	0	1	$\frac{35}{105}$	$\frac{70}{105}$	1	$\frac{35}{105}$	$\frac{70}{105}$	1	1
$Q_{2,4}$	0	0	$\frac{45}{105}$	0	$\frac{60}{105}$	$\frac{45}{105}$	0	$\frac{60}{105}$	$\frac{45}{105}$	1	$\frac{60}{105}$	$\frac{45}{105}$	1	$\frac{60}{105}$	1	1
Q_0	0	$\frac{36}{420}$	$\frac{80}{420}$	$\frac{160}{420}$	$\frac{144}{420}$	$\frac{116}{420}$	$\frac{196}{420}$	$\frac{180}{420}$	$\frac{240}{420}$	$\frac{224}{420}$	$\frac{304}{420}$	$\frac{276}{420}$	$\frac{260}{420}$	$\frac{340}{420}$	$\frac{384}{420}$	1

Finally, for $\epsilon = \frac{1}{2}$, define $\widehat{Bel}_\epsilon = \frac{1}{2}Q_0 + \frac{1}{2}\widehat{Bel}$, whose values are reported below

$\mathcal{F}$	$\emptyset$	1	2	3	4	12	13	14	23	24	34	123	124	134	234	Ω
$\widehat{Bel}_\epsilon$	0	$\frac{60}{420}$	$\frac{40}{420}$	$\frac{80}{420}$	$\frac{72}{420}$	$\frac{100}{420}$	$\frac{140}{420}$	$\frac{132}{420}$	$\frac{120}{420}$	$\frac{112}{420}$	$\frac{272}{420}$	$\frac{180}{420}$	$\frac{172}{420}$	$\frac{332}{420}$	$\frac{360}{420}$	1
$\widehat{\mu}_\epsilon$	0	$\frac{60}{420}$	$\frac{40}{420}$	$\frac{80}{420}$	$\frac{72}{420}$	0	0	0	0	0	$\frac{120}{420}$	0	0	0	$\frac{48}{420}$	0

We have that $\widehat{Bel}_\epsilon$ is an equivalent inner approximating one-step Choquet martingale belief function for $\underline{Q}$, furthermore, it is the lower envelope of the class of probability measures on $\mathcal{F}$

$$\widehat{\mathcal{Q}}_\epsilon = \{Q \in \mathbf{P}(\Omega, \mathcal{F}) \,:\, Q = (1-\epsilon)Q_0 + \epsilon Q', Q' \in \mathrm{core}(\widehat{Bel})\}.$$

A direct computation shows that

$$\mathbb{C}_{\widehat{Bel}_\epsilon}\left(\frac{S_1^1}{S_0^1}\right) = \min_{Q \in \widehat{\mathcal{Q}}_\epsilon} \mathbb{E}_Q\left(\frac{S_1^1}{S_0^1}\right) = 1 + r.$$

♦

The use of d_1 has been justified in Miranda et al. (2021, 2022); Montes et al. (2018, 2019) since this distance is the most intuitive as it measures the imprecision added when we replace $\underline{Q}$ with Bel. Despite using d_1 we get a linear programming problem the main disadvantage is that the optimal solution is generally not unique, as shown in the following example.

Example 4.11 *Consider $\Omega = \{1,2,3,4\}$, $m_1 = 5$, $m_2 = 3$, $m_3 = 2$, $m_4 = \frac{1}{2}$ and $1 + r = 4$. According to Theorem 4.1, we have $I = \{1\}$, $J = \{2,3,4\}$ and* $\mathrm{ext}(\mathrm{cl}(\mathcal{Q})) = \{Q_{1,2}, Q_{1,3}, Q_{1,4}\}$ *inducing $\underline{Q}$ and μ reported below*

$\mathcal{F}$	$\emptyset$	1	2	3	4	12	13	14	23	24	34	123	124	134	234	Ω
$Q_{1,2}$	0	$\frac{9}{18}$	$\frac{9}{18}$	0	0	1	$\frac{9}{18}$	$\frac{9}{18}$	$\frac{9}{18}$	$\frac{9}{18}$	0	1	1	$\frac{9}{18}$	$\frac{9}{18}$	1
$Q_{1,3}$	0	$\frac{12}{18}$	0	$\frac{6}{18}$	0	$\frac{12}{18}$	1	$\frac{12}{18}$	$\frac{6}{18}$	0	$\frac{6}{18}$	1	$\frac{12}{18}$	1	$\frac{6}{18}$	1
$Q_{1,4}$	0	$\frac{14}{18}$	0	0	$\frac{4}{18}$	$\frac{14}{18}$	$\frac{14}{18}$	1	0	$\frac{4}{18}$	$\frac{4}{18}$	$\frac{14}{18}$	1	1	$\frac{4}{18}$	1
$\underline{Q}$	0	$\frac{9}{18}$	0	0	0	$\frac{12}{18}$	$\frac{9}{18}$	$\frac{9}{18}$	0	0	0	$\frac{14}{18}$	$\frac{12}{18}$	$\frac{9}{18}$	$\frac{4}{18}$	1
μ	0	$\frac{9}{18}$	0	0	0	$\frac{3}{18}$	0	0	0	0	0	$\frac{2}{18}$	0	0	$\frac{4}{18}$	0

The following two belief functions have Möbius inverse minimizing the distance d_1

$\mathcal{F}$	$\emptyset$	1	2	3	4	12	13	14	23	24	34	123	124	134	234	Ω
$\widehat{\mu}_1$	0	$\frac{12}{18}$	0	0	0	0	0	0	$\frac{4}{18}$	0	0	$\frac{2}{18}$	0	0	0	0
$\widehat{Bel}_1$	0	$\frac{12}{18}$	0	0	0	$\frac{12}{18}$	$\frac{12}{18}$	$\frac{12}{18}$	$\frac{4}{18}$	0	0	1	$\frac{12}{18}$	$\frac{12}{18}$	$\frac{4}{18}$	1
$\widehat{\mu}_2$	0	$\frac{11}{18}$	0	0	0	$\frac{3}{18}$	0	0	$\frac{4}{18}$	0	0	0	0	0	0	0
$\widehat{Bel}_2$	0	$\frac{11}{18}$	0	0	0	$\frac{14}{18}$	$\frac{11}{18}$	$\frac{11}{18}$	$\frac{4}{18}$	0	0	1	$\frac{14}{18}$	$\frac{11}{18}$	$\frac{4}{18}$	1

and it holds that $d_1(\widehat{Bel}_1, \underline{Q}) = d_1(\widehat{Bel}_2, \underline{Q}) = \frac{20}{18}$.

♦

The problem of not uniqueness of the solution could be dealt with some further procedures adopting different criteria as, for example, taking into account measures of specificity (Yager, 1983) or other distances (Miranda et al., 2021) in order to select a unique solution among those that are optimal according to d_1.

The main feature of distance d_2 is that problem (4.23) admits a unique optimal solution as the objective function turns out to be strictly convex (see, e.g., Montes et al., 2018, 2019; Miranda et al., 2022). In other terms, the choice of d_2 amounts in computing the orthogonal projection of $\underline{Q}$ onto the set of inner approximating one-step Choquet martingale belief functions for $\underline{Q}$.

Example 4.12 *Consider Ω, m_1, m_2, $\underline{m_2}$, m_4, and $1+r$ as in Example 4.11. In this case, the unique optimal solution $\widehat{Bel}$ minimizing d_2 has Möbius inverse $\widehat{\mu}$ such that $\widehat{\mu}(1) = 0.628655$, $\widehat{\mu}(2) = 0.0087719$, $\widehat{\mu}(12) = 0.149123$, $\widehat{\mu}(23) = 0.18421$, $\widehat{\mu}(234) = 0.0292399$, and 0 otherwise.*

In this case we have $d_2(\widehat{Bel}, \underline{Q}) = 0.169591$.

♦

Distances d_1 and d_2 face the problem of approximation from a metric point of view. Another possibility is to take as d the *Bregman divergence* induced by a bounded *(strictly) proper scoring rule* (see, e.g., Censor and Zenios, 1997; Predd et al., 2009). Indeed, as shown in Petturiti and Vantaggi (2023), proper scoring rules allow to introduce a notion of coherence for belief functions through a penalty criterion that generalizes classical results for probabilities (see, e.g., Predd et al., 2009). In particular, as shown in Gilio and Sanfilippo (2011), every bounded proper scoring rule gives rise to a Bregman divergence that can be used in the approximation. It actually turns out that d_2 is the Bregman divergence induced by the Brier quadratic scoring rule, so it has a justification in terms of the penalty criterion for belief functions (see Petturiti and Vantaggi, 2023). We stress that if we take a Bregman divergence for d, then (4.23) is generally a non-linear problem with linear constraints.

Furthermore, besides minimizing a distance or a divergence, other approaches are available, like minimizing a measure of non-specificity (or imprecision) as done in Denœux (2006).

Finally, we consider the Čebišëv distance, denoted by d_∞, defined as

$$d_\infty(Bel_1, Bel_2) = \max_{A \in \mathcal{F}} |Bel_1(A) - Bel_2(A)|. \tag{4.25}$$

The optimization problem (4.23), that is equivalent to (4.22), taking into account that $\widehat{Bel} \geq \underline{Q}$, can be written in terms of Möbius inverse as

$$d_\infty(\widehat{Bel}, \underline{Q}) = \max_{A \in \mathcal{F}} (\widehat{Bel}(A) - \underline{Q}(A)) = \max_{A \in \mathcal{U}} \left(\sum_{\emptyset \neq B \subseteq A} \widehat{\mu}(B) - \underline{Q}(A) \right). \tag{4.26}$$

Since $\underline{Q}(A)$, for every $A \in \mathcal{U}$, is constant, the objective function is a maximum of linear (affine) functions, i.e., it is a continuous piecewise-linear function. In this case, problem (4.23) can be transformed in an equivalent linear programming problem by adding a new variable t and a linear constraint for every component in the maximum of the objective function

$$
\text{minimize } t
$$

$$
\text{subject to:}
$$

$$
\begin{cases}
\sum_{\emptyset \neq B \subseteq A} \widehat{\mu}(B) - \underline{Q}(A) \leq t, & \text{for every } A \in \mathcal{U}, \\
\sum_{\emptyset \neq B \subseteq A} \widehat{\mu}(B) \geq \underline{Q}(A), & \text{for every } A \in \mathcal{U}, \\
\sum_{i=1}^{n} m_i \left(\sum_{\{i\} \subseteq B \subseteq \{1,\ldots,i\}} \widehat{\mu}(B) \right) = 1 + r, & \\
\sum_{B \in \mathcal{U}} \widehat{\mu}(B) = 1, & \\
\widehat{\mu}(B) \geq 0, & \text{for every } B \in \mathcal{U}.
\end{cases} \tag{4.27}
$$

As occurs with d_1, distance d_∞ does not give, generally, a unique solution to problem (4.27).

The following example shows that the inner approximating one-step Choquet martingale belief function that minimizes distance d_∞ is not unique. Further, the optimal solutions can fail to be undominated.

Example 4.13 *Consider* $\Omega = \{1,2,3,4\}$, $m_1 = 4$, $m_2 = 2$, $m_3 = \frac{1}{2}$, $m_4 = \frac{1}{4}$, $1 + r = 1$, *and* $\mathcal{Q}$ *and* $\underline{Q}$ *of Example 4.1. Its lower envelope* $\underline{Q}$ *has two (and so it has infinitely many) inner approximating one-step Choquet martingale belief functions* $\widehat{Bel}_1, \widehat{Bel}_2$ *minimizing the distance* d_∞*:*

$\mathcal{F}$	$\emptyset$	1	2	3	4	12	13	14	23	24	34	123	124	134	234	Ω
$\underline{Q}$	0	0	0	0	0	$\frac{15}{105}$	0	0	0	0	$\frac{60}{105}$	$\frac{21}{105}$	$\frac{15}{105}$	$\frac{60}{105}$	$\frac{84}{105}$	1
μ	0	0	0	0	0	$\frac{15}{105}$	0	0	0	0	$\frac{60}{105}$	$\frac{6}{105}$	0	0	$\frac{24}{105}$	0
$\widehat{\mu}_1$	0	$\frac{272}{2100}$	$\frac{152}{2100}$	0	0	$\frac{148}{2100}$	0	0	$\frac{120}{2100}$	0	$\frac{1200}{2100}$	0	0	0	$\frac{208}{2100}$	0
$\widehat{Bel}_1$	0	$\frac{272}{2100}$	$\frac{152}{2100}$	0	0	$\frac{572}{2100}$	$\frac{272}{2100}$	$\frac{272}{2100}$	$\frac{272}{2100}$	$\frac{152}{2100}$	$\frac{1200}{2100}$	$\frac{692}{2100}$	$\frac{572}{2100}$	$\frac{1472}{2100}$	$\frac{1680}{2100}$	1
$\widehat{\mu}_2$	0	$\frac{272}{2100}$	$\frac{181}{2100}$	0	0	$\frac{119}{2100}$	0	0	$\frac{91}{2100}$	0	$\frac{1200}{2100}$	$\frac{29}{2100}$	0	0	$\frac{208}{2100}$	0
$\widehat{Bel}_2$	0	$\frac{272}{2100}$	$\frac{181}{2100}$	0	0	$\frac{572}{2100}$	$\frac{272}{2100}$	$\frac{272}{2100}$	$\frac{272}{2100}$	$\frac{181}{2100}$	$\frac{1200}{2100}$	$\frac{692}{2100}$	$\frac{572}{2100}$	$\frac{1472}{2100}$	$\frac{1680}{2100}$	1

for which we have that $d_\infty(\widehat{Bel}_1, \underline{Q}) = d_\infty(\widehat{Bel}_2, \underline{Q}) = \frac{272}{2100}$. *Moreover, we have that* $\widehat{Bel}_1 \neq \widehat{Bel}_2$ *but* $\underline{Q} \leq \widehat{Bel}_1 \leq \widehat{Bel}_2$, *thus the* d_∞*-optimal solution* $\widehat{Bel}_2$ *is dominated by the* d_∞*-optimal solution* $\widehat{Bel}_1$.

♦

Up to now, we have considered only the lower price assessment $\underline{\pi}'(S_1^1) = S_0^1$. Nevertheless, we also know that $\max_{Q \in \mathrm{cl}(\mathcal{Q})} (1+r)^{-1} \mathbb{E}_Q(S_1^1) = S_0^1$, thus we could consider an upper price assessment for the stock. If we further impose to respect the upper price assessment $\overline{\pi}'(S_1^1) = S_0^1$, then the notion of one-step Choquet martingale belief function given in Definition 4.2 can be strengthened as follows.

Definition 4.3 *A belief function* $\widehat{Bel}$ *on* $\mathcal{F}$ *is called:*

- *a one-step* **strong Choquet martingale belief function** *if*

$$(1+r)^{-1}\mathbb{C}_{\widehat{Bel}}(S_1^1) = S_0^1 \quad \textit{and} \quad (1+r)^{-1}\mathbb{C}_{\widehat{Bel}}(-S_1^1) = -S_0^1;$$

- *an* **inner approximating one-step strong Choquet martingale belief function** *for $\underline{Q}$ if it is both an inner approximation for $\underline{Q}$ and a one-step strong Choquet martingale belief function;*

- *an* **equivalent inner approximating one-step strong Choquet martingale belief function** *for $\underline{Q}$ if it is an inner approximating one-step strong Choquet martingale belief function for $\underline{Q}$ and $\widehat{Bel} \sim P$.*

Still referring to a distance d defined on $\mathbf{B}(\Omega, \mathcal{F})$, an optimal inner approximating one-step strong Choquet martingale belief function $\widehat{Bel}$ for $\underline{Q}$ can be found by solving the following optimization problem:

$$\text{minimize } d(\widehat{Bel}, \underline{Q})$$
$$\text{subject to:}$$
$$\begin{cases} \widehat{Bel}(A) \geq \underline{Q}(A), & \text{for every } A \in \mathcal{F}, \\ (1+r)^{-1}\mathbb{C}_{\widehat{Bel}}(S_1^1) = S_0^1, & \\ (1+r)^{-1}\mathbb{C}_{\widehat{Bel}}(-S_1^1) = -S_0^1, & \\ \widehat{Bel} \in \mathbf{B}(\Omega, \mathcal{F}). & \end{cases} \tag{4.28}$$

Also in this case, problem (4.28) can be reformulated as follows

$$\text{minimize } d(\widehat{Bel}, \underline{Q})$$
$$\text{subject to:}$$
$$\begin{cases} \sum\limits_{\emptyset \neq B \subseteq A} \widehat{\mu}(B) \geq \underline{Q}(A), & \text{for every } A \in \mathcal{U}, \\ \sum\limits_{i=1}^{n} m_i \left(\sum\limits_{\{i\} \subseteq B \subseteq \{1,\ldots,i\}} \widehat{\mu}(B) \right) = 1+r, & \\ \sum\limits_{i=1}^{n} m_i \left(\sum\limits_{\{i\} \subseteq B \subseteq \{i,\ldots,n\}} \widehat{\mu}(B) \right) = 1+r, & \\ \sum\limits_{B \in \mathcal{U}} \widehat{\mu}(B) = 1, & \\ \widehat{\mu}(B) \geq 0, & \text{for every } B \in \mathcal{U}. \end{cases} \tag{4.29}$$

The following theorem states that any inner approximating one-step strong Choquet martingale belief function $\widehat{Bel}$ for $\underline{Q}$ is actually a probability measure belonging to $\mathrm{cl}(\mathcal{Q})$.

Theorem 4.6
For every distance d defined on $\mathbf{B}(\Omega, \mathcal{F})$, the set of feasible solutions of problem (4.28) is $\mathrm{cl}(\mathcal{Q})$. Further, if $d = d_1$ then the set of optimal solutions of problem (4.28) coincides with $\mathrm{cl}(\mathcal{Q})$, while if $d = d_2$ then there is a unique optimal solution.

Proof. First notice that inner approximating one-step strong Choquet martingale belief functions for $\underline{Q}$, that is feasible solutions of problem (4.28), are in one-to-one correspondence with feasible solutions of problem (4.29). Define $\mathcal{V} = \mathcal{U} \setminus \{\{i\} : i \in \Omega\}$. Subtracting memberwise the second equation to the third equation of problem (4.29) we get

$$\sum_{B \in \mathcal{V}} \left(\max_{i \in B} m_i - \min_{i \in B} m_i \right) \widehat{\mu}(B) = 0.$$

Hence, since for every $B \in \mathcal{V}$, $\left(\max_{i \in B} m_i - \min_{i \in B} m_i \right) > 0$ and $\widehat{\mu}(B) \geq 0$, any feasible solution of problem (4.29) is such that $\widehat{\mu}(B) = 0$, for every $B \in \mathcal{V}$. In turn, this implies that any feasible solution of problem (4.29) is the Möbius inverse of a probability measure which is a feasible solution of problem (4.28). Thus, if $\widehat{Bel}$ is a feasible solution of problem (4.28) we have $\mathbb{C}_{\widehat{Bel}}\left(\frac{S_1^1}{S_0^1}\right) = \mathbb{E}_{\widehat{Bel}}\left(\frac{S_1^1}{S_0^1}\right) = 1 + r$, implying that $\widehat{Bel} \in \mathrm{cl}(\mathcal{Q})$. Vice versa, every element of $\mathrm{cl}(\mathcal{Q})$ is easily seen to be a feasible solution of problem (4.28).

If $d = d_1$, since every feasible solution $\widehat{Bel}$ of problem (4.28) is a probability measure with Möbius inverse $\widehat{\mu}$, by (4.24) we get that

$$d_1(\widehat{Bel}, \underline{Q}) = \sum_{\{i\} \in \mathcal{U}} 2^{|\Omega \setminus \{i\}|} \widehat{\mu}(\{i\}) - \sum_{A \in \mathcal{U}} \underline{Q}(A) = 2^{n-1} - \sum_{A \in \mathcal{U}} \underline{Q}(A), \tag{4.30}$$

that does not depend on $\widehat{Bel}$. Hence, all the elements of $\mathrm{cl}(\mathcal{Q})$ are optimal according to d_1.

If $d = d_2$, then the uniqueness of the optimal solution immediately follows since the objective function of (4.28) is strictly convex.

□

Hence, using d_1 any element of $\mathrm{cl}(\mathcal{Q})$ turns out to be optimal, while using d_2 we get the orthogonal projection of $\underline{Q}$ onto the set of inner approximating one-step strong Choquet martingale belief functions for $\underline{Q}$, which is, by Theorem 4.6, the set $\mathrm{cl}(\mathcal{Q})$.

Example 4.14 *Consider Ω, m_1, m_2, m_3, m_4, and $1+r$ as in Example 4.11. Using d_1, the set of optimal inner approximating strong Choquet martingale belief functions for $\underline{Q}$ is $\mathrm{cl}(\mathcal{Q})$ and for every $\widehat{Bel} \in \mathrm{cl}(\mathcal{Q})$ we have $d_1(\widehat{Bel}, \underline{Q}) = \frac{8}{3}$.*

On the other hand, if we use the distance d_2 we have a unique optimal solution which is the following inner approximating one-step strong Choquet martingale belief function (probability measure)

Ω	1	2	3	4
$\widehat{Bel}$	0.638889	0.188596	0.102339	0.0701755

for which we have $d_2(\widehat{Bel}, \underline{Q}) = 0.572124$.

♦

Let us stress that, for a fixed $Q_0 \in \mathcal{Q}$, if $\widehat{Bel}$ is an inner approximating one-step strong Choquet martingale belief function (probability measure), then

$\mathrm{core}(\widehat{Bel}) = \{\widehat{Bel}\}$. In this case $\widehat{\mathcal{Q}}_\epsilon$ reduces to the singleton $\widehat{\mathcal{Q}}_\epsilon = \{(1-\epsilon)Q_0 + \epsilon\widehat{Bel}\}$, thus the lower envelope $\underline{Q}_\epsilon$ is an equivalent martingale measure, that is $\underline{Q}_\epsilon = (1-\epsilon)Q_0 + \epsilon\widehat{Bel} \in \mathcal{Q}$.

The following proposition states that, both for d_1 and d_2, an optimal inner approximating one-step martingale (strong martingale) belief function does not dominate any other inner approximating one-step martingale (strong martingale) belief function. This is inline with results proved in Miranda et al. (2022); Montes et al. (2018, 2019).

Proposition 4.1
Let $d = d_1$ or $d = d_2$. If $\widehat{Bel}$ is an optimal solution of problem (4.22) (problem (4.28)) then there is no feasible solution $\widehat{Bel}'$ of problem (4.22) (problem (4.28)) such that $\widehat{Bel}' \neq \widehat{Bel}$ and $\underline{Q} \leq \widehat{Bel}' \leq \widehat{Bel}$.

Proof. The proof can be obtained by a straightforward adaptation of Lemma 14 in Montes et al. (2018).

□

If $d = d_\infty$, problem (4.29) can be turned into the following equivalent linear programming problem

$$\text{minimize } t$$

$$\text{subject to:}$$

$$\begin{cases} \sum_{\emptyset\neq B\subseteq A} \widehat{\mu}(B) - \underline{Q}(A) \leq t, & \text{for every } A \in \mathcal{U}, \\ \sum_{\emptyset\neq B\subseteq A} \widehat{\mu}(B) \geq \underline{Q}(A), & \text{for every } A \in \mathcal{U}, \\ \sum_{i=1}^{n} m_i \left(\sum_{\{i\}\subseteq B\subseteq\{1,\dots,i\}} \widehat{\mu}(B) \right) = 1 + r, & \\ \sum_{i=1}^{n} m_i \left(\sum_{\{i\}\subseteq B\subseteq\{i,\dots,n\}} \widehat{\mu}(B) \right) = 1 + r, & \\ \sum_{B\in\mathcal{U}} \widehat{\mu}(B) = 1, & \\ \widehat{\mu}(B) \geq 0, & \text{for every } B \in \mathcal{U}. \end{cases} \tag{4.31}$$

By Theorem 4.6, all feasible solution of problem (4.31) are Möbius inverses of probability measures which are equivalent martingale measures. In particular, every such $\widehat{\mu}$ is zero on $\mathcal{V} = \mathcal{U} \setminus \{\{i\} : i \in \Omega\}$. This implies that

$$d_\infty(\widehat{Bel}, \underline{Q}) = \max_{A\in\mathcal{F}} (\widehat{Bel}(A) - \underline{Q}(A)) = \max_{A\in\mathcal{U}} \left(\sum_{i\in A} \widehat{\mu}(\{i\}) - \underline{Q}(A) \right). \tag{4.32}$$

Chapter 5

A multi-period binomial pricing model under Dempster-Shafer uncertainty

The one-period n-nomial market model characterized in Chapter 4 leads us to think to a model that can allow frictions in the pricing process where the (lower) pricing rule is defined as a (discounted) Choquet expectation. However it does not allow to model a multi-period process.

Our aim is to extend the one-period model under Dempster-Shafer uncertainty to a multi-period binomial random process, taking the framework of belief functions as our natural environment to model uncertainty.

As stressed in the previous Chapter, this framework allows us to incorporate naturally frictions that are present in real markets.

In Section 5.1 we define a binomial random process in a multi-period setting with respect to a belief function that can be interpreted as a *lower* price process.

In order to generalize the binomial model with respect to a (additive) probability measure (in Section 2.2.1) into an *imprecise* binomial model such that it is mathematically tractable, we ask to satisfy a suitable version of Markov and time-homogeneity properties with respect to belief functions, through the *product (geometric)* conditioning rule (see Section 1.2.3). A *global* belief function that satisfies the desired properties and that is also mathematically interpretable is proved to exist and it is characterized through its set of k-step *transition belief functions*, since the usual Chapman-Kolmogorov equation does not apply (see Section 1.1). The transition belief functions are completely determined by the choice of two parameters such that, if they sum up to one, the process collapses into the classical multiplicative binomial process.

In order to define a (lower) pricing rule by means of the conditional Choquet expectation operator, the belief function has to satisfy the martingale property, as proposed in the one-period setting in Definition 4.2.

Our notion of *Dempster-Shafer multiplicative binomial process* differs from other proposals of *imprecise* Markov process, summarized in Section 3.2, since they focus on local models rather than on a global one and they work with interval probabilities or capacities.

The random process we characterize allows us to introduce, in Sections 5.2 and 5.3, a dynamic lower pricing rule expressed by a (discounted) conditional Choquet expectation operator with a closed form expression that is a *one-step* Choquet martingale but it is only a *global* Choquet *super-martingale*, i.e., when more than one-step is considered.

In Section 5.4, we propose a dynamic (lower) pricing rule for a European type derivative whose payoff depends only on the lower price of a stock whose (lower) price process is a Dempster-Shafer multiplicative binomial process. Then, we need to provide a normative justification of the proposed dynamic lower pricing rule by referring to a dynamic generalized no-arbitrage condition that is the extension to the multi-period setting of Theorems 4.4–4.5 in Chapter 4, continuing to be under *partially resolving uncertainty.*

5.1 Dempster-Shafer multiplicative binomial processes

Our goal is the extension of the one-period model in Chapter 4 to a multi-period setting, taking the framework of belief functions as our natural environment, with the hope that the generalized no-arbitrage principle presented in Section 4.3 could be suitable in this novel setting. Another argument in favor of belief functions, as underlined in Chapter 1, is the interpretability of the belief function as a measure of evidence that other non-additive measures, such as 2-monotone capacities, may not have (see note p. 25).

As in the previous Chapters, we continue to denote a belief function by Bel, its Möbius inverse by μ. We refer to the *product (or geometric) conditioning rule* for belief functions, presented in Section 1.2.3 (in order to overcome heavy notation, we denote it without the apex ^{P}): for every $E, H \in \mathcal{F}$ with $Bel(H) > 0$

$$Bel(E|H) = \frac{Bel(E \cap H)}{Bel(H)}. \tag{5.1}$$

The above rule of conditioning can be reformulated as a *chain rule* for belief functions: for every $E, H \in \mathcal{F}$ with $Bel(H) > 0$ it holds that

$$Bel(E \cap H) = Bel(E|H)Bel(H). \tag{5.2}$$

Remark 5. Whenever $Bel(H) > 0$, $Bel(\cdot|H)$ is a belief function on $\mathcal{F}$ inducing $\mathrm{core}(Bel(\cdot|H))$. Thus, locally on every H with positive belief, $Bel(\cdot|H)$ can be interpreted as a coherent lower probability on $\mathcal{F}$. Nevertheless, setting $\mathcal{H} = \{H \in \mathcal{F} : Bel(H) > 0\}$, the function $Bel(\cdot|\cdot)$ on $\mathcal{G} = \mathcal{F} \times \mathcal{H}$ may fail to be a coherent lower conditional probability in the sense of Williams (Williams, 2007).

Note 16: *The choice of the product (geometric) conditioning rule, in addition to the greater simplicity of calculation, is motivated by the fact that relations between conditioning rules (see* (1.42)–(1.43)*) imply that, for all* $X \in \mathbb{R}^\Omega$*,*

$$\oint X(i)\,\mathrm{d}Bel^G(i|H) \leq \min\left\{\oint X(i)\,\mathrm{d}Bel(i|H), \oint X(i)\,\mathrm{d}Bel^D(i|H)\right\}. \tag{5.3}$$

Although the product (geometric) conditioning rule cannot be interpreted as the lower envelope of conditional probabilities computed with respect to $\mathrm{core}(Bel)$*, it produces* *less dilation with respect to the general (Bayes) conditioning rule. It has an important consequence on bid-ask pricing since, for a fixed unconditional belief function, we generally obtain narrower bid-ask price intervals.*

Given $Bel(\cdot|H)$ on $\mathcal{F}$, then it uniquely extends to a conditional completely monotone functional defined on $\mathbb{R}^\Omega$ through the Choquet integral by setting, for all $X \in \mathbb{R}^\Omega$,

$$\oint X(i)\,\mathrm{d}Bel(i|H) = \sum_{i=1}^{n} (X(\sigma(i)) - X(\sigma(i+1)))\, Bel(E_i^\sigma|H), \tag{5.4}$$

where σ is a permutation of Ω as defined in Section 1.2.2. As in the unconditional setting (see *(Ch.8)*, p.19), the above Choquet integral can be given a lower expectation interpretation locally on H, by referring to core$(Bel(\cdot|H))$, as it holds that

$$\oint X(i)\mathrm{d}Bel(i|H) = \min_{P \in \mathrm{core}(Bel(\cdot|H))} \int X(i)\mathrm{d}P(i), \tag{5.5}$$

where the integrals in the minimum are of Stieltjes type.

Consider a discrete time finite-horizon stochastic process $\{X_0, \ldots, X_T\}$ with $T \in \mathbb{N}$, $X_0 = x_0 > 0$ and, for $t = 1, \ldots, T$, that is the analogous of the binomial risky asset's price process in (2.22)

$$X_t = \begin{cases} uX_{t-1} & \text{if "up"}, \\ dX_{t-1} & \text{if "down"}, \end{cases} \tag{5.6}$$

where $u > d > 0$ are the "up" and "down" coefficients. In what follows, such a process will be called a *multiplicative binomial process*. The process above is defined on a filtered measurable space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t=0}^T)$, where $\Omega = \{1, \ldots, 2^T\}$ and $\mathcal{F}_t$ is the algebra generated by random variables $\{X_0, \ldots, X_t\}$, for $t = 0, \ldots, T$, with $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and $\mathcal{F}_T = \mathcal{F} = \mathcal{P}(\Omega)$.

The trajectories of $\{X_0, \ldots, X_T\}$ can be represented graphically on a binomial tree. In particular, every state $i \in \Omega$ is identified with the path corresponding to the T-digit binary expansion of number $i - 1$, in which zeroes are interpreted as "up" movements and ones as "down" movements.

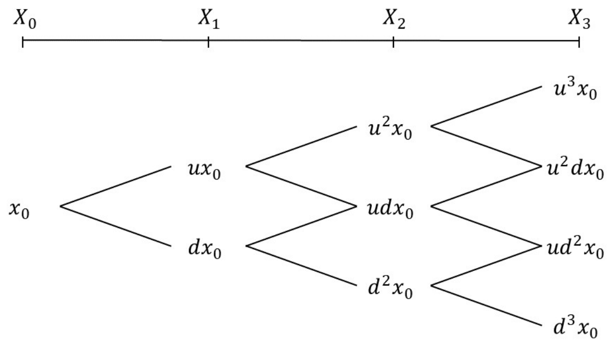

Figure 5.1. Binomial tree for $T = 3$.

Assume that uncertainty on the evolution of the process is not additive but is handled in the Dempster-Shafer theory of evidence through a belief function

$Bel : \mathcal{F} \to [0,1]$. From now on, we assume there is a filtered belief space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t=0}^T, Bel)$ with a fixed belief function Bel.

For $t = 1, \ldots, T$, denote

$$\mathcal{A}_t = \{a_j = u^j d^{t-j} : j = 0, \ldots, t\}, \tag{5.7}$$

for which we have $a_0 < a_1 < \ldots < a_t$ and, for $a_i \leq a_h$, let

$$[a_i, a_h] = \{a_j \in \mathcal{A}_t : a_i \leq a_j \leq a_h\}. \tag{5.8}$$

For every $x > 0$ and $A \in \mathcal{P}(\mathcal{A}_t)$, denote

$$Ax = \{a_j x : a_j \in A\}, \tag{5.9}$$

where $Ax = \emptyset$ if $A = \emptyset$. In particular, each random variable X_t takes values x_t in $\mathcal{X}_t = \mathcal{A}_t x_0$.

Definition 5.1 *Given a filtered belief space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t=0}^T, Bel)$, the process $\{X_0, \ldots, X_T\}$ is said to satisfy the:*

Markov property: *if for every $0 \leq t \leq T-1$ and $1 \leq k \leq T-t$, $A \in \mathcal{P}(\mathcal{A}_k)$, and $x_0 \in \mathcal{X}_0, \ldots, x_t \in \mathcal{X}_t$ on a trajectory with positive belief it holds that*

$$Bel(X_{t+k} \in Ax_t | X_0 = x_0, \ldots, X_t = x_t) = Bel(X_{t+k} \in Ax_t | X_t = x_t);$$

time-homogeneity property: *if for every $0 \leq t \leq T-1$ and $1 \leq k \leq T-t$, $A \in \mathcal{P}(\mathcal{A}_k)$, and $x_0 \in \mathcal{X}_0, \ldots, x_t \in \mathcal{X}_t$ on a trajectory with positive belief it holds that*

$$Bel(X_{t+k} \in Ax_t | X_0 = x_0, \ldots, X_t = x_t) = \beta_k(A),$$

where $\beta_k : \mathcal{P}(\mathcal{A}_k) \to [0,1]$ is a fixed belief function.

If the process satisfies both the properties above is called a **DS-multiplicative binomial process** *(where DS reads "Dempster-Shafer").*

In the particular case where the process $\{X_0, \ldots, X_T\}$ satisfies the Markov property, then the time-homogeneity property reduces to

$$Bel(X_{t+k} \in Ax_t | X_t = x_t) = \beta_k(A). \tag{5.10}$$

The properties above are called *one-step* if they hold only for $k = 1$.

The first issue to face is the existence of a belief function Bel on $\mathcal{F}$ that makes the process $\{X_0, \ldots, X_T\}$ Markov and time-homogeneous (i.e., a DS-multiplicative binomial process).

Note 17: *A DS-multiplicative binomial process singles out a family of belief functions $\{\beta_k : k = 1, \ldots, T\}$ defined on the family of power sets $\{\mathcal{P}(\mathcal{A}_k) : k = 1, \ldots, T\}$ that, in turn, are determined by the particular Bel that is chosen. If Bel is not additive, then we need the entire family of β_k's since the usual Chapman-Kolmogorov equation (see Section 1.1) does not apply due to the lack of additivity. Such β_k's are actually* **k-step transition belief functions**.

Let b_u, b_d be two strictly positive parameters with $b_u + b_d \leq 1$ that are intuitively interpreted as one-step "up" and "down" conditional beliefs.

As discussed in Cinfrignini et al. (2022) related to a DS-*additive* binomial process, the first idea is to recover the probabilistic construction by imposing an analogue of (2.23) for the terminal value: for every trajectory on the binomial tree with $x_0 \in \mathcal{X}_0, \ldots, x_T \in \mathcal{X}_T$ and $x_T = u^j d^{T-j} x_0$, with $j = 0, \ldots, T$, we have

$$Bel(X_0 = x_0, \ldots, X_T = x_T) = b_u^j b_d^{T-j}. \tag{5.11}$$

Since the event $\{X_0 = x_0, \ldots, X_T = x_T\} \in \mathcal{X}_T$ corresponding to a trajectory reduces to a singleton $i \in \Omega$, then (5.11) implies the following Möbius inverse

$$\mu(X_0 = x_0, \ldots, X_T = x_T) = b_u^j b_d^{T-j}. \tag{5.12}$$

The first immediate consequence is that Bel satisfying (5.12) is positive on $\mathcal{F} \setminus \{\emptyset\}$, hence conditioning through the product rule is always well-defined.

Assumption (5.11) provides just few constraints for the Bel since, given that the sum over all the possible trajectories of the binomial tree is

$$\sum_{x_0,\ldots,x_T} \mu(X_0 = x_0, \ldots, X_T = x_T) = (b_u + b_d)^T, \tag{5.13}$$

if $b_u + b_d < 1$, we need to allocate the remaining mass of $(1 - (b_u + b_d)^T)$.

Between the infinite number of belief functions satisfying (5.11) there are some failing both Markov and time-homogeneity properties and some others just failing one of them.

Proposition 5.1
Let Bel be the belief function on $\mathcal{F}$ whose Möbius inverse satisfies (5.12) and such that $\mu(\Omega) = 1 - (b_u + b_d)^T$. Then the process $\{X_0, \ldots, X_T\}$ satisfies the Markov property.

Proof. It can be straightforwardly proven considering the proof of Proposition 1 in Cinfrignini et al. (2022) with respect to a multiplicative binomial process instead of an additive binomial process.

□

However, the belief function defined as in Proposition 5.1 fails the time-homogeneity property, as shown in the following example adapted from Cinfrignini et al. (2022).

Example 5.1 *Let $T = 3$ and consider the belief function Bel defined in Proposition 5.1. Simple computations show that*

$$Bel(X_3 = ud^2x_0 | X_0 = x_0, X_1 = ux_0, X_2 = udx_0) = \frac{b_d}{b_u + b_d},$$

$$Bel(X_1 = dx_0 | X_0 = x_0) = b_d(b_u + b_d)^2.$$

The (one-step) time-homogeneity property is not satisfied.

♦

Moreover, the belief function introduced in Proposition 5.1 does not respect the intuitive meaning of parameters b_u and b_d, since we have, for every $0 < t \leq T-1$, that

$$Bel(X_{t+1} = ux_t | X_0 = x_0, \ldots, X_t = x_t) = \frac{b_u}{b_u + b_d} \tag{5.14}$$

$$Bel(X_{t+1} = dx_t | X_0 = x_0, \ldots, X_t = x_t) = \frac{b_d}{b_u + b_d}. \tag{5.15}$$

Hence, we directly impose the one-step time-homogeneity requiring Bel to satisfy, for every $0 \leq t \leq T-1$ and $x_0 \in \mathcal{X}_0, \ldots, x_t \in \mathcal{X}_t$ on a trajectory with positive belief

$$Bel(X_{t+1} = ux_t | X_0 = x_0, \ldots, X_t = x_t) = b_u, \tag{5.16}$$

$$Bel(X_{t+1} = dx_t | X_0 = x_0, \ldots, X_t = x_t) = b_d. \tag{5.17}$$

Note 18: *A straightforward application of the chain rule* (5.2) *shows that* (5.16)–(5.17) *imply* (5.11)*. Hence, imposing* (5.16)–(5.17) *produces much more constraints for the belief function Bel.*

The following proposition (adapted from Cinfrignini et al., 2022) proves that there exists (at least) one belief function Bel satisfying constraints (5.16)–(5.17).

Proposition 5.2
A belief function Bel on $\mathcal{F}$ satisfies (5.16)–(5.17) (that is, *one-step time-homogeneity*) if and only if the corresponding Möbius inverse μ satisfies the following conditions:

(i) for $x_T = u^j d^{T-j} x_0 \in \mathcal{X}_T$ with $j = 0, \ldots, T$

$$\mu(\{X_0 = x_0, \ldots, X_T = x_T\}) = b_u^j b_d^{T-j}; \tag{5.18}$$

(ii) for $x_{T-1} = u^j d^{T-1-j} x_0 \in \mathcal{X}_{T-1}$ with $j = 0, \ldots, T-1$

$$\mu(\{X_0 = x_0, \ldots, X_{T-1} = x_{T-1}\}) = b_u^j b_d^{T-1-j}(1 - b_u - b_d); \tag{5.19}$$

(iii) for every $0 < t < T-1$, for $x_t = u^j d^{t-j} x_0 \in \mathcal{X}_t$,

$$\sum_{\substack{B \subseteq \{X_0 = x_0, \ldots, X_t = x_t\} \\ B \not\subseteq \{X_0 = x_0, \ldots, X_{t+1} = ux_t\} \\ B \not\subseteq \{X_0 = x_0, \ldots, X_{t+1} = dx_t\}}} \mu(B) = b_u^j b_d^{t-j}(1 - b_u - b_d); \tag{5.20}$$

(iv)

$$\sum_{\substack{B \subseteq \{X_0 = x_0\} \\ B \not\subseteq \{X_0 = x_0, X_1 = ux_0\} \\ B \not\subseteq \{X_0 = x_0, X_1 = dx_0\}}} \mu(B) = (1 - b_u - b_d); \tag{5.21}$$

where all events $\{X_0 = x_0, \ldots, X_t = x_t\}$ correspond to partial trajectories on the binomial tree.

Proof. It can be straightforwardly proven from Proposition 2 in Cinfrignini et al. (2022).

□

However, the following example shows that *Bel* as defined in Proposition 5.2, that satisfies the (one-step) time-homogeneity property, does not imply time-homogeneity or Markov property.

Example 5.2 *Let* $T = 3$ *and* $\Omega = \{1, 2, 3, 4, 5, 6, 7, 8\}$ *with*

Ω	X_0	X_1	X_2	X_3
1	x_0	ux_0	u^2x_0	u^3x_0
2	x_0	ux_0	u^2x_0	u^2dx_0
3	x_0	ux_0	udx_0	u^2dx_0
4	x_0	ux_0	udx_0	ud^2x_0
5	x_0	dx_0	udx_0	u^2dx_0
6	x_0	dx_0	udx_0	ud^2x_0
7	x_0	dx_0	d^2x_0	ud^2x_0
8	x_0	dx_0	d^2x_0	d^3x_0

Consider the Möbius inverse μ *on* $\mathcal{F}$ *satisfying conditions (i),(ii) of Proposition 5.2, and*

$$\mu(\{2,3\}) = b_u(1-(b_u+b_d)), \qquad \mu(\{5,6,7,8\}) = b_d(1-(b_u+b_d)),$$
$$\mu(\{3,5\}) = \mu(\Omega) = \frac{1}{2}(1-(b_u+b_d)),$$

while it is zero otherwise.

Such μ *satisfies conditions (iii) and (iv) of Proposition 5.2, thus the corresponding Bel satisfy the one-step time-homogeneity property* (5.16)–(5.17)*.*

However, such Bel does not satisfy the Markov property, as

$$Bel(X_3 = u^2dx_0 | X_0 = x_0, X_1 = ux_0, X_2 = udx_0) = b_u,$$
$$Bel(X_3 = u^2dx_0 | X_2 = udx_0) = \frac{2b_u^2b_d + \frac{1}{2}(1-(b_u+b_d))}{2b_ub_d + \frac{1}{2}(1-(b_u+b_d))}.$$

Such Bel does not satisfy time-homogeneity for $k > 1$ *since, for* $a_1 = ud \in \mathcal{A}_2$*, we have*

$$Bel(X_3 = ua_1x_0 | X_0 = x_0, X_1 = ux_0) = 2b_ub_d + (1-(b_u+b_d)),$$
$$Bel(X_2 = a_1x_0 | X_0 = x_0) = 2b_ub_d + \frac{1}{2}(1-(b_u+b_d)).$$

In turn, let $b_u = b_d = \delta \in (0, \frac{1}{2})$ *and consider the Möbius inverse* $\tilde{\mu}$ *on* $\mathcal{F}$ *satisfying conditions (i),(ii) of Proposition 5.2 and*

$$\widetilde{\mu}(\{1,2,3,4\}) = \widetilde{\mu}(\{5,6,7,8\}) = \widetilde{\mu}(\{3,5\}) = \widetilde{\mu}(\{4,6\}) = \delta(1-2\delta),$$
$$\widetilde{\mu}(\{3,4,5,6\}) = (1-2\delta)^2,$$

and it is 0 *otherwise. It is easily verified that also conditions (iii) and (iv) of Proposition 5.2 are satisfied, thus the corresponding* $\widetilde{Bel}$ *satisfies the one-step time-homogeneity. We notice that*

$$\widetilde{Bel}(X_3 = x_3 | X_0 = x_0, X_1 = x_1, X_2 = x_2) = \widetilde{Bel}(X_3 = x_3 | X_2 = x_2) = \delta,$$
$$\widetilde{Bel}(X_3 \in Ax_1 | X_0 = x_0, X_1 = x_1) = \widetilde{Bel}(X_3 \in Ax_1 | X_1 = x_1),$$

for every $A \in \mathcal{P}(\mathcal{A}_2)$ and $x_t \in \mathcal{P}(\mathcal{A}_t)$, with $t = 0, 1, 2, 3$, hence, $\widetilde{Bel}$ satisfies the Markov property. However, for $a_1 = ud \in \mathcal{A}_2$ we have

$$\widetilde{Bel}(X_3 = ua_1x_0 | X_0 = x_0, X_1 = ux_0) = 2\delta^2,$$
$$\widetilde{Bel}(X_2 = a_1x_0 | X_0 = x_0) = 2\delta^2 + 1 - 2\delta.$$

Hence the time-homogeneity property does not hold.

♦

In general, we can have infinitely many belief functions on $\mathcal{F}$ that make the process $\{X_0, \ldots, X_T\}$ a DS-multiplicative binomial process as the following example shows.

Example 5.3 *Let $T = 3$ and $\Omega = \{1, \ldots, 8\}$ with the same identification of states as Example 5.2. Let $\alpha \in [0, 1]$ and consider the Möbius inverse μ_α on $\mathcal{F}$ satisfying conditions (i) and (ii) of Proposition 5.2 and such that, denoting $h = (1 - b_u - b_d)$:*

$$\mu_\alpha(\{1,2,3,4\}) = b_u h, \quad \mu_\alpha(\{5,6,7,8\}) = b_d h,$$
$$\mu_\alpha(\{1,2,3,4,5,6,7\}) = \alpha h,$$
$$\mu_\alpha(\{2,3,4,5,6,7,8\}) = (1-\alpha)h.$$

It is easily shown that the corresponding belief function Bel is a DS-multiplicative binomial process. The family of belief functions $\{\beta_1^\alpha, \beta_2^\alpha, \beta_3^\alpha\}$, with $\mathcal{A}_t$ defined in (5.7) and $A^t_{i_1,\ldots,i_m} = \{a^t_{i_1}, \ldots, a^t_{i_m}\}$, is

$\mathcal{P}(\mathcal{A}_1)$	$\emptyset$	A_0^1	A_1^1	$\mathcal{A}_1$
β_1^α	0	b_d	b_u	1

$\mathcal{P}(\mathcal{A}_2)$	$\emptyset$	A_0^2	A_1^2	A_2^2	$A_{0,1}^2$	$A_{0,2}^2$	$A_{1,2}^2$	$\mathcal{A}_2$
β_2^α	0	b_d^2	$2b_ub_d$	b_u^2	$b_ub_d + b_d$	$b_u^2 + b_d^2$	$b_ub_d + b_u$	1

$\mathcal{P}(\mathcal{A}_3)$	$\emptyset$	A_0^3	A_1^3	A_2^3
β_3^α	0	b_d^3	$3b_ub_d^2$	$3b_u^2b_d$
$\mathcal{P}(\mathcal{A}_3)$	A_3^3	$A_{0,1}^3$	$A_{0,2}^3$	$A_{0,3}^3$
β_3^α	b_u^3	$2b_ub_d^2 + b_d^2$	$b_d^3 + 3b_u^2b_d$	$b_d^3 + b_u^3$
$\mathcal{P}(\mathcal{A}_3)$	$A_{1,2}^3$	$A_{1,3}^3$	$A_{2,3}^3$	$A_{0,1,2}^3$
β_3^α	$b_ub_d^2 + b_u^2b_d + 2b_ub_d$	$3b_ub_d^2 + b_u^3$	$2b_u^2b_d + b_u^2$	$b_d(b_u^2 + b_u + 1) + (1-\alpha)h$
$\mathcal{P}(\mathcal{A}_3)$	$A_{0,1,3}^3$	$A_{0,2,3}^3$	$A_{1,2,3}^3$	$\mathcal{A}_3$
β_3^α	$2b_ub_d^2 + b_d^2 + b_u^3$	$2b_u^2b_d + b_u^2 + b_d^3$	$b_u(b_d^2 + b_d + 1) + \alpha h$	1

Therefore we have an infinite class of belief functions $\{Bel_\alpha : \alpha \in [0, 1]\}$ that make the process a DS-multiplicative binomial process.

♦

Actually, some choices of Bel on $\mathcal{F}$ could lead to a lack of interpretation for the family $\{\beta_k : k = 1, \ldots, T\}$ induced by Bel, and to a large amount of parameters that could make difficult a calibration procedure. This is why, in what follows, we restrict to a particular family of k-step transition belief functions that guarantee a clear interpretation and a nice parametrization.

We continue to require that, for every $0 \leq t \leq T-1$, and $x_0 \in \mathcal{X}_0, \ldots, x_t \in \mathcal{X}_t$ on a trajectory with positive belief, the belief function Bel satisfies the one-step time-homogeneity (5.16)–(5.17), since, in this way, b_u and b_d are interpreted as one-step "up" and "down" conditional beliefs.

Remark 6. In case $b_u + b_d = 1$, conditions (5.16)–(5.17) determine a unique additive belief function (probability measure) that satisfies time-homogeneity and Markov properties.

Assuming that $b_u + b_d < 1$, we need to characterize Bel by means of the k-step transition belief functions β_k's.

Assumption 1. *We endow $\mathcal{F}$ with a belief function Bel whose Möbius inverse μ is such that*

(a) for $x_T = u^j d^{T-j} x_0 \in \mathcal{X}_T$ with $j = 0, \ldots, T$

$$\mu(\{X_0 = x_0, \ldots, X_T = x_T\}) = b_u^j b_d^{T-j}; \tag{5.22}$$

(b) for $0 < t < T$ and $x_t = u^j d^{t-j} x_0 \in \mathcal{X}_t$ with $j = 0, \ldots, t$

$$\mu(\{X_0 = x_0, \ldots, X_t = x_t\}) = b_u^j b_d^{t-j}(1 - (b_u + b_d)); \tag{5.23}$$

(c)

$$\mu(\{X_0 = x_0\}) = \mu(\Omega) = 1 - (b_u + b_d); \tag{5.24}$$

(d) μ is zero otherwise.

Our aim is to derive the conditional belief distributions of random variables X_t's. For that, we consider the belief function $\beta_k : \mathcal{P}(\mathcal{A}_k) \to [0,1]$ that is interpreted as the k-step *transition belief functions*. Consider that the belief function β_k is characterized by the following Möbius inverse $\mu'_k : \mathcal{P}(\mathcal{A}_k) \to [0,1]$:

(1) for $j = 0, \ldots, k$

$$\mu'_k(\{a_j\}) = \binom{k}{j} b_u^j b_d^{k-j}; \tag{5.25}$$

(2) for $h = 1, \ldots, k$ and $j = 0, \ldots, k-h$

$$\mu'_k([a_j, a_{j+h}]) = \binom{k-h}{j} b_u^j b_d^{k-h-j}(1 - (b_u + b_d)); \tag{5.26}$$

(3) μ'_k is zero otherwise.

The following proposition shows that μ'_k is indeed the Möbius inverse of a belief function β_k on $\mathcal{P}(\mathcal{A}_k)$.

Proposition 5.3
Let $\mu'_k : \mathcal{P}(\mathcal{A}_k) \to [0,1]$ be defined as in (5.25)–(5.26) and it is zero otherwise, then μ'_k is the Möbius inverse of a belief function $\beta_k : \mathcal{P}(\mathcal{A}_k) \to [0,1]$ defined, for all $A \in \mathcal{A}_k$ as

$$\beta_k(A) = \sum_{a_j \in A} \binom{k}{j} b_u^j b_d^{k-j} + \sum_{\substack{[a_j, a_{j+h}] \subseteq A \\ h \geq 1}} \binom{k-h}{j} b_u^j b_d^{k-h-j} (1 - (b_u + b_d)). \quad (5.27)$$

Proof. First we need to show that μ'_k is the Möbius inverse of a belief function. The function μ'_k is easily seen to be non-negative, moreover it sums up to 1 since

$$\sum_{j=0}^{k} \mu'_k(\{a_j\}) + \sum_{h=1}^{k} \sum_{j=0}^{k-h} \mu'_k([a_j, a_{j+h}]) =$$
$$= (b_u + b_d)^k + \sum_{h=1}^{k} (b_u + b_d)^{k-h} (1 - (b_u + b_d)) = 1.$$

Finally, the claim follows since, for all $A \in \mathcal{P}(\mathcal{A}_k)$, we have that

$$\beta_k(A) = \sum_{a_j \in A} \mu'(\{a_j\}) + \sum_{\substack{[a_j, a_{j+h}] \subseteq A \\ h \geq 1}} \mu'([a_j, a_{j+h}]),$$

that is μ'_k is the Möbius inverse of the belief function β_k.

□

Note 19: *Notice that* β_k *in* (5.27) *is consistent with* (5.16)–(5.17) *(one-step time-homogeneity) as it holds that* $\beta_1(\emptyset) = 0$, $\beta_1(\{u\}) = b_u$, $\beta_1(\{d\}) = b_d$, *and* $\beta_1(\mathcal{A}_1) = 1$. *This leads to a clear interpretation where ambiguity that amounts to the excessive weight to unity* $(1 - (b_u + b_d))$ *is attached to the entire frame of evidence* $\mathcal{A}_1 = \{d, u\}$.

The belief function β_k in (5.27) generalizes the binomial distribution with parameters k and b_u, to which it reduces in case $b_u + b_d = 1$, since the second summation vanishes. On the other hand, if $b_u + b_d < 1$, then the second summation takes into account a contribution of intervals contained in A which receive a binomial-like weighting deflated by the excessive weight to unity $(1 - (b_u + b_d))$. More in detail, we have that intervals of length h contribute by weights mimicking the binomial distribution with parameters $k - h$ and b_u, multiplied by the deflator $(1 - (b_u + b_d))$. Looking at the binomial tree representation of process $\{X_0, \ldots, X_T\}$ we get that, starting from a node x_t at time t and looking ahead of k steps, the interval $[a_j, a_{j+h}]$ of length h represents the set of all trajectories starting at node x_t and continuing for k steps that have a fixed state x_{k-h} at time $k - h$. Indeed, all the continuations of partial trajectory $x_t, \ldots, x_{k+h}$ for the remaining h times will end in a state belonging to $[a_j, a_{j-h}]x_t$. Therefore, interpreting such weights as *evidence* in the spirit of Dempster-Shafer theory (Shafer, 1976a), $\beta_k(A)$ is obtained by summing the binomial-like weights of all partial trajectories with decreasing length starting from node x_t, that support the evidence of having a final state of the process after k steps belonging to Ax_t.

The following theorem states that the belief function Bel whose Möbius inverse is given in Assumption 1 meets all the desiderata.

Theorem 5.1
Let $Bel : \mathcal{F} \to [0,1]$ be characterized by a Möbius inverse $\mu : \mathcal{F} \to [0,1]$ as defined in Assumption 1. Then, a multiplicative binomial process on the filtered belief space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t=0}^T, Bel)$ meets the following properties:

(i) $Bel(B) > 0$, for every $B \in \mathcal{F} \setminus \{\emptyset\}$;

(ii) $\{X_0, \ldots, X_T\}$ is a DS-multiplicative binomial process whose transition belief functions $\{\beta_k : k = 1, \ldots, T\}$ satisfy (5.27).

Proof. We assume $\mu : \mathcal{F} \to [0,1]$ as defined is Assumption 1. We prove statement *(i)*. We have that $\mu(B) \geq 0$, for all $B \in \mathcal{F}$, moreover

- $\sum_{x_T \in \mathcal{X}_T} \mu(\{X_0 = x_0, \ldots, X_T = x_T\}) = (b_u + b_d)^T$,
- for all $0 < t < T$, $\sum_{x_t \in \mathcal{X}_t} \mu(\{X_0 = x_0, \ldots, X_t = x_t\}) = (b_u + b_d)^t (1 - (b_u + b_d))$,
- $\mu(\{X_0 = x_0\}) = \mu(\Omega) = 1 - (b_u + b_d)$,

while μ is zero otherwise. Hence, we get that

$$\sum_{B \in \mathcal{F}} \mu(B) = (b_u + b_d)^T + \sum_{0<t<T} (b_u + b_d)^t (1 - (b_u + b_d)) + (1 - (b_u + b_d)) = 1,$$

that is μ is the Möbius inverse of a belief function. Moreover, since elements of Ω can be identified with the trajectories on the binomial tree, i.e., with events $\{X_0 = x_0, \ldots, X_T = x_T\}$, μ is such that $\mu(\{i\}) > 0$, for all $i \in \Omega$. In turn, this implies that Bel is such that $Bel(B) > 0$, for every $B \in \mathcal{F} \setminus \{\emptyset\}$.

We prove statement *(ii)*. For every $0 \leq t \leq T$, we let $x_t = u^j d^{t-j} x_0 \in \mathcal{X}_t$ and prove that

$$Bel(X_0 = x_0, X_1 = x_1, \ldots, X_t = u^j d^{t-j} x_0) = b_u^j b_d^{t-j}. \tag{5.28}$$

In order to get the events with strictly positive μ contained in the event $\{X_0 = x_0, X_1 = x_1, \ldots, X_t = u^j d^{t-j} x_0\}$, the corresponding partial trajectory on the binomial tree must be completed for the remaining $T - t$ times indexed by l with $l = T - t, T - t - 1, \ldots, 0$, working backward.

For $l = T - t$ we have to add $i_{T-t} = 0, \ldots, T - t$ movements to the state of the random variable X_t. For a fixed i_{T-t}, by summing over all the possible completions of the trajectory, we have that

$$\sum_{x_{t+1}, \ldots, x_{T-1}} \mu(X_0 = x_0, \ldots, X_t = u^j d^{t-j} x_0, \ldots, X_T = u^{j+i_{T-t}} d^{t-j+(T-t)-i_{T-t}} x_0) =$$
$$= \binom{T-t}{i_{T-t}} b_u^{j+i_{T-t}} b_d^{k-j+(T-t)-i_{T-t}}.$$

Then, summing over i_{T-t} we have that

$$\sum_{i_{T-t}=0}^{T-t} \binom{T-t}{i_{T-t}} b_u^{j+i_{T-t}} b_d^{t-j+(T-t)-i_{T-t}}. \tag{5.29}$$

For a generic $0 \leq l \leq T-t-1$ we need to add $i_l = 0, \dots, l$ movements to the state of the random variable X_t. For a fixed i_l, by summing over all the possible completions of the trajectory, we have that

$$\begin{aligned}
&\sum_{x_{t+1},\dots,x_{t+l-1}} \mu(X_0 = x_0, \dots, X_t = u^j d^{t-j} x_0, \dots, X_{t+l} = u^{j+i_l} d^{t-j+l-i_l} x_0) = \\
&= \binom{l}{i_l} b_u^{j+i_l} b_d^{t-j+l-i_l} (1 - (b_u + b_d)).
\end{aligned}$$

Then, summing over i_l we have that

$$\sum_{i_l=0}^{l} \binom{l}{i_l} b_u^{j+i_l} b_d^{t-j+l-i_l} (1 - (b_u + b_d)). \tag{5.30}$$

Therefore we obtain that

$$\begin{aligned}
&Bel(X_0 = x_0, X_1 = x_1, \dots, X_t = u^j d^{t-j} x_0) = \\
&= \sum_{i_{T-t}=0}^{T-t} \binom{T-t}{i_{T-t}} b_u^{j+i_{T-t}} b_d^{t-j+(T-t)-i_{T-t}} + \\
&\quad + \sum_{l=0}^{T-t-1} \sum_{i_l=0}^{l} \binom{l}{i_l} b_u^{j+i_l} b_d^{t-j+l-i_l} (1 - (b_u + b_d)) = \\
&= b_u^j b_d^{t-j} \left[(b_u + b_d)^{T-t} + (1 - (b_u + b_d)) \sum_{l=0}^{T-t-1} (b_u + b_d)^l \right] = \\
&= b_u^j b_d^{t-j}.
\end{aligned}$$

Now we prove that

$$Bel(X_t = u^j d^{t-j} x_0) = \binom{t}{j} b_u^t b_d^{t-j}. \tag{5.31}$$

(5.29) considers the trajectory from time t to time T, having fixed the part before t. Summing over all the possible completions of the trajectory before time t, we get

$$\binom{t}{j} \sum_{i_{T-t}=0}^{T-t} \binom{T-t}{i_{T-t}} b_u^{j+i_{T-t}} b_d^{t-j+(T-t)-i_{T-t}}.$$

Analogously, for a generic $0 \leq l \leq T-t-1$, (5.30) considers the trajectory from time t to time $t+l$, having fixed the part before t. For a fixed l, summing over all the possible completions of the trajectory before time t, we get

$$\binom{t}{j} \sum_{i_l=0}^{l} \binom{l}{i_l} b_u^{j+i_l} b_d^{t-j+l-i_l} (1 - (b_u + b_d)).$$

Hence, we obtain

$$
\begin{aligned}
&Bel(X_t = u^j d^{t-j} x_0) = \\
&\quad = \binom{t}{j} \sum_{i_{T-t}=0}^{T-t} \binom{T-t}{i_{T-t}} b_u^{j+i_{T-t}} b_d^{t-j+(T-t)-i_{T-t}} + \\
&\quad\quad + \sum_{l=0}^{T-t-1} \binom{t}{j} \sum_{i_l=0}^{l} \binom{l}{i_l} b_u^{j+i_l} b_d^{t-j+l-i_l} (1-(b_u+b_d)) = \\
&\quad = \binom{t}{j} b_u^j b_d^{t-j}.
\end{aligned}
$$

Now let $1 \leq k \leq T-t$ and $A \subseteq \mathcal{A}_k = \{a_z = u^z d^{k-z} \,:\, z = 0, \ldots, k\}$. Let μ'_k be the Möbius inverse of β_k defined in Proposition 5.3 through (5.25)–(5.26). We prove that

$$
Bel(\{X_{t+k} \in A u^j d^{t-j} x_0\} \cap \{X_0 = x_0, \ldots, X_t = u^j d^{t-j} x_0\}) = b_u^j b_d^{t-j} \beta_k(A). \quad (5.32)
$$

If $a_z \in A$, summing over the partial trajectories from time $t+1$ to time $t+k-1$, we get that

$$
\begin{aligned}
&\sum_{x_{t+1},\ldots,x_{t+k-1}} \mu(X_0 = x_0, \ldots, X_t = u^j d^{t-j} x_0, \ldots, X_{t+k} = u^{z+j} d^{t+k-(z+j)} x_0) = \\
&\quad = \binom{k}{z} b_u^{z+j} b_d^{t+k-(z+j)} = \\
&\quad = b_u^j b_d^{t-j} \mu'_k(\{a_z\}).
\end{aligned}
$$

If $h \geq 1$ and $[a_z, a_{z+h}] \subseteq A$, summing over the partial trajectories from time $t+1$ to time $t+k-h-1$, we get that

$$
\begin{aligned}
&\sum_{x_{t+1},\ldots,x_{t+k-h-1}} \mu(X_0 = x_0, \ldots, X_t = u^j d^{t-j} x_0, \ldots, X_{t+k-h} = u^{z+j} d^{t+k-h-(z+j)} x_0) = \\
&\quad = \binom{k-h}{z} b_u^{z+j} b_d^{t+k-h-(z+j)} (1-(b_u+b_d)) = \\
&\quad = b_u^j b_d^{t-j} \mu'_k([a_z, a_{z+h}]).
\end{aligned}
$$

Hence

$$\begin{aligned}
&Bel(\{X_{t+k} \in Au^j d^{t-j} x_0\} \cap \{X_0 = x_0, \ldots, X_t = u^j d^{t-j} x_0\}) = \\
&= \sum_{a_z \in A} \binom{k}{z} b_u^{z+j} b_d^{t+k-z-j} + \\
&\quad + \sum_{\substack{[a_j, a_{j+h}] \subseteq A \\ h \geq 1}} \binom{k-h}{z} b_u^{z+j} b_d^{t+k-h-z-j} (1 - (b_u + b_d)) = \\
&= b_u^j b_d^{t-j} \left[\sum_{a_z \in A} \mu'_k(\{a_z\}) + \sum_{\substack{[a_j, a_{j+h}] \subseteq A \\ h \geq 1}} \mu'_k([a_z, a_{z+h}]) \right] = \\
&= b_u^j b_d^{t-j} \beta_k(A).
\end{aligned}$$

Proceeding in analogy with the derivation of (5.31) we get that

$$Bel(\{X_{t+k} \in Au^j d^{t-j} x_0\} \cap \{X_t = u^j d^{t-j} x_0\}) = \binom{t}{j} b_u^j b_d^{t-j} \beta_k(A). \tag{5.33}$$

Finally, Markovianity and time-homogeneity follow from (5.25)–(5.26), (5.31)–(5.33) since we obtain

$$\begin{aligned}
&Bel(X_{t+k} \in Au^j d^{t-j} x_0 | X_0 = x_0, \ldots, X_t = u^j d^{t-j} x_0) = \\
&\quad = Bel(X_{t+k} \in Au^j d^{t-j} x_0 | X_t = u^j d^{t-j} x_0) \\
&\quad = \beta_k(A).
\end{aligned}$$

□

We summarize the proposed belief functions' structure and their properties.

μ	*one-step t.h.*	*t.h.*	*Markov*	*Interpretation*
(5.12) $\mu(\Omega)=1-(b_u+b_d)^T$	×	×	✓	*no*
(5.18)–(5.21)	✓	×	×	b_u and b_d are one-step "up" and "down" conditional beliefs
(5.22)–(5.24)	✓	✓	✓	β_k is a generalization of the binomial distribution

Assumption 2. *From now on, we assume the belief function Bel meeting conditions (i)–(ii) of Theorem 5.1 to be fixed. Therefore, we always refer to transition belief functions* $\{\beta_k : k = 1, \ldots, T\}$ *satisfying* (5.27).

Every DS-multiplicative binomial process can be associated with an additive binomial process through a logarithmic transformation. In detail, we consider the process $\{R_0, \ldots, R_T\}$ where

$$R_t = \ln \frac{X_t}{X_0}, \quad \text{for } t = 0, \ldots, T. \tag{5.34}$$

Setting $l_u = \ln u$ and $l_d = \ln d$, we have that $R_0 = 0$ and R_t ranges in the set

$$\mathcal{R}_t = \{r_j = jl_u + (t-j)l_d \, : \, j = 0, \ldots, t\}. \tag{5.35}$$

The process $\{R_0, \ldots, R_T\}$ is still a time-homogeneous Markov process under Bel, since it satisfies:

(Markovianity) For every $0 \leq t \leq T-1$ and $1 \leq k \leq T-t$, and $B \in \mathcal{P}(\mathcal{R}_k)$,

$$Bel(R_{t+k} \in B | R_0 = 0, \ldots, R_t = r_t) = Bel(R_{t+k} \in B | R_t = r_t).$$

(Time-homogeneity) For every $0 \leq t \leq T-1$ and $1 \leq k \leq T-t$, and $B \in \mathcal{P}(\mathcal{R}_k)$,

$$Bel(R_{t+k} \in B | R_t = r_t) = \beta_k(\exp(B)).$$

Remark 7. From a financial point of view, if $\{X_0, \ldots, X_T\}$ is used to model the price evolution of a stock, then $\{R_0, \ldots, R_T\}$ is the corresponding log-return process. We also notice that $\{R_0, \ldots, R_T\}$ is an example of *DS-random walk* as introduced in Cinfrignini et al. (2022).

We denote, for every $0 \leq t \leq T$, a random variable as $Y : \Omega \to \mathbb{R}$ that, as usual (see Section 1.1) is said to be $\mathcal{F}_t$*-measurable* if it is constant on the atoms of the (natural) algebra $\mathcal{F}_t$ generated by $\{X_0, \ldots, X_t\}$. Notice that, all random variables $Y \in \mathbb{R}^\Omega$ are $\mathcal{F}_T$-measurable, since $\mathcal{F}_T = \mathcal{F} = \mathcal{P}(\Omega)$

Definition 5.2 *Let $\{X_0, \ldots, X_T\}$ be a DS-multiplicative binomial process on the filtered belief space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t=0}^T, Bel)$. Then, for every random variable $Y \in \mathbb{R}^\Omega$, define*

$$\begin{aligned}
\mathbb{C}[Y|X_t = x_t] &= \oint Y(i) Bel(\mathrm{d}i | X_t = x_t), \\
\mathbb{C}[Y|X_0 = x_0, \ldots, X_t = x_t] &= \oint Y(i) Bel(\mathrm{d}i | X_0 = x_0, \ldots, X_t = x_t).
\end{aligned}$$

In turn, we define the random variables $\mathbb{C}[Y|X_t]$ and $\mathbb{C}[Y|X_0, \ldots, X_t]$ setting, for all $i \in \{X_t = x_t\}$,

$$\mathbb{C}[Y|X_t](i) := \mathbb{C}[Y|X_t = x_t], \tag{5.36}$$

and, for all $i \in \{X_0 = x_0, \ldots, X_t = x_t\}$,

$$\mathbb{C}[Y|X_0, \ldots, X_t](i) := \mathbb{C}[Y|X_0 = x_0, \ldots, X_t = x_t]. \tag{5.37}$$

We also simply write

$$\mathbb{C}[Y|\mathcal{F}_t] := \mathbb{C}[Y|X_0, \ldots, X_t], \tag{5.38}$$

which is easily seen to be $\mathcal{F}_t$-measurable. The operator $\mathbb{C}[\cdot|\mathcal{F}_t]$ will be referred to as *conditional Choquet expectation*, in the rest of the Chapter. The Choquet integral with respect to a belief function implies that $\mathbb{C}[\cdot|\mathcal{F}_t]$ satisfies properties *(Ch.3)*, *(Ch.6)*, *(Ch.9)*, *(Ch.8)* (p.19), and it further satisfies the following property.

Proposition 5.4
The conditional Choquet expectation $\mathbb{C}[\cdot|\mathcal{F}_t]$ associated with the filtered belief space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t=0}^T, Bel)$ satisfies:

(conditional constant) for all $\mathcal{F}_t$-measurable $Y \in \mathbb{R}^\Omega$ and all $Z \in \mathbb{R}^\Omega$,

$$\mathbb{C}[Y|\mathcal{F}_t] = Y,$$

and, if $Y \geq 0$,

$$\mathbb{C}[YZ|\mathcal{F}_t] = Y\mathbb{C}[Z|\mathcal{F}_t].$$

Proof. The property is an immediate consequence of properties of the Choquet integral with respect to a belief function, see Section 1.2.2, and with respect to (5.36)–(5.37).

□

Remark 8. It is easy to verify that the conditional Choquet expectation $\mathbb{C}[\cdot|\mathcal{F}_t]$ may fail to satisfy the *tower property* (see (1.8), p.9), that is, in general, for $0 \leq t \leq T-1$ and $1 \leq k \leq T-t$, we have that

$$\mathbb{C}[\mathbb{C}[Y|\mathcal{F}_{t+k}]|\mathcal{F}_t] \neq \mathbb{C}[Y|\mathcal{F}_t].$$

If $\varphi(x)$ is a real-valued function of one real variable defined on the range of X_{t+k}, then the following proposition characterizes the conditional Choquet expectation when $Y = \varphi(X_{t+k})$.

Proposition 5.5
Let $\{X_0, \ldots, X_T\}$ be a DS-multiplicative binomial process on the filtered belief space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t=0}^T, Bel)$. Then, for every $0 \leq t \leq T-1$ and $1 \leq k \leq T-t$, and every real-valued function of one real variable $\varphi(x)$ defined on the range of X_{t+k}, we have that

$$\begin{aligned}\mathbb{C}[\varphi(X_{t+k})|X_t = x_t] &= \sum_{z=0}^{k} \varphi(a_z x_t)\binom{k}{z} b_u^z b_d^{k-z} \\ &+ \sum_{h=1}^{k}\sum_{z=0}^{k-h}\left[\min_{a_i\in[a_z,a_{z+h}]}\varphi(a_i x_t)\right]\binom{k-h}{z} b_u^z b_d^{k-h-z}(1-(b_u+b_d))\end{aligned}$$

and $\mathbb{C}[\varphi(X_{t+k})|X_0 = x_0, \ldots, X_t = x_t] = \mathbb{C}[\varphi(X_{t+k})|X_t = x_t]$.

In particular,

(i) if $\varphi(x)$ is non-decreasing

$$\begin{aligned}\mathbb{C}[\varphi(X_{t+k})|X_t = x_t] &= \sum_{z=0}^{k} \varphi(u^z d^{k-z} x_t)\binom{k}{z} b_u^z b_d^{k-z} \\ &+ \sum_{z=0}^{k-1}\varphi(u^z d^{k-z} x_t)\sum_{h=1}^{k-z}\binom{k-h}{z} b_u^z b_d^{k-h-z}(1-(b_u+b_d));\end{aligned}$$

(ii) if $\varphi(x)$ is non-increasing

$$\begin{aligned}\mathbb{C}[\varphi(X_{t+k})|X_t = x_t] &= \sum_{z=0}^{k} \varphi(u^z d^{k-z} x_t) \binom{k}{z} b_u^z b_d^{k-z} \\ &+ \sum_{z=0}^{k-1} \varphi(u^{k-z} d^z x_t) \sum_{h=1}^{k-z} \binom{k-h}{z} b_u^{k-h-z} b_d^z (1 - (b_u + b_d));\end{aligned}$$

(iii) in the particular case $Y = \varphi(X_{t+k}) = X_{t+k}$

$$\begin{aligned}\mathbb{C}[X_{t+k}|X_t = x_t] &= \sum_{z=0}^{k} u^z d^{k-z} x_t \binom{k}{z} b_u^z b_d^{k-z} \\ &+ \sum_{z=0}^{k-1} u^z d^{k-z} x_t \sum_{h=1}^{k-z} \binom{k-h}{z} b_u^z b_d^{k-h-z} (1 - (b_u + b_d)),\end{aligned}$$

and $\mathbb{C}[X_{t+k}|X_0 = x_0, \ldots, X_t = x_t] = \mathbb{C}[X_{t+k}|X_t = x_t]$.

Proof. Conditionally on $\{X_t = x_t\}$, the random variable X_{t+k} takes values in $\mathcal{A}_k x_t$ and has belief distribution given by β_k on $\mathcal{P}(\mathcal{A}_k)$. Let μ'_k be the Möbius inverse of β_k defined in Proposition 5.3 through (5.25)–(5.26). The general expression of $\mathbb{C}[\varphi(X_{t+k})|X_t = x_t]$ easily follows by the properties of the Choquet integral. We have that

$$\begin{aligned}\mathbb{C}[\varphi(X_{t+k})|X_t = x_t] &= \oint \varphi(X_{t+k}(i)) Bel(\mathrm{d}i|X_t = x_t) = \\ &= \oint_{\mathcal{A}_k} \varphi(a x_t) \beta_k(\mathrm{d}a) = \\ &= \sum_{z=0}^{k} \varphi(a_z x_t) \mu'_k(\{a_z\}) \\ &\quad + \sum_{h=1}^{k} \sum_{z=0}^{k-h} \left[\min_{a_i \in [a_z, a_{z+h}]} \varphi(a_i x_t) \right] \mu'_k([a_z, a_{z+h}]),\end{aligned}$$

and the claim follows by (5.25)–(5.26). The equality $\mathbb{C}[\varphi(X_{t+k})|X_0 = x_0, \ldots, X_t = x_t] = \mathbb{C}[\varphi(X_{t+k})|X_t = x_t]$ follows by the time-homogeneity and Markov properties of the process. The cases of a non-decreasing or non-increasing $\varphi(x)$ are obtained by computing minima and gathering terms. Finally, the particular case $\varphi(X_{t+k}) = X_{t+k}$ is obtained

$$\begin{aligned}
\mathbb{C}[X_{t+k}|X_t = x_t] &= \oint X_{t+k}(i) Bel(\mathrm{d}i|X_t = x_t) = \\
&= \oint_{\mathcal{A}_k} a x_t \beta_k(\mathrm{d}a) = \\
&= \sum_{z=0}^{k} a_z x_t \mu'_k(\{a_z\}) + \sum_{h=1}^{k}\sum_{z=0}^{k-h} a_z x_t \mu'_k([a_z, a_{z+h}]) = \\
&= \sum_{z=0}^{k} u^z d^{k-z} x_t \binom{k}{z} b_u^z b_d^{k-z} \\
&\quad + \sum_{z=0}^{k-1} u^z d^{k-z} x_t \sum_{h=1}^{k-z} \binom{k-h}{z} b_u^z b_d^{k-h-z}(1-(b_u+b_d)).
\end{aligned}$$

Also, in this particular case, the equality $\mathbb{C}[X_{t+k}|X_0 = x_0, \ldots, X_t = x_t] = \mathbb{C}[X_{t+k}|X_t = x_t]$ follows by the time-homogeneity and Markov properties of the process.

□

Note 20: *Our DS-multiplicative binomial process differs from other proposals of imprecise Markov process, summarized in Section 3.2, basically not (only) for the assumption of time-homogeneity but because they generally characterize a local model (i.e., for one-step time-interval) and extend it to a multi-step model, or do not work in the belief functions setting, which makes them unsuitable for achieving our desired properties. In particular, T'Joens et al. (2021); Krak et al. (2019); Nendel (2021) work with lower/upper expectations without characterizing a family of transition non-additive measures. Also, Kast et al. (2014) characterize a symmetric random walk with respect to a constant conditional capacity and dynamic consistency is axiomatically obtained, while Škulj (2009, 2006) considers random walks and does not assume time-homogeneity. Finally, the binomial (probability) distribution is a particular case of our DS-multiplicative binomial process, that does not hold for all the quoted approaches.*

5.2 Equivalent Choquet martingale belief functions

We consider a financial market composed by two assets: a risk-free bond and a risky stock that does not pay dividends. Contrary to the classical binomial pricing model in Section 2.2.1, we allow frictions in the market in the form of bid-ask spreads. Therefore, every security is characterized by a *lower* and an *upper price* and we will focus in modelling the lower prices evolution. We assume that the evolution of the lower prices of the two securities is described by the processes

$$\left\{\underline{S}_0^0, \ldots, \underline{S}_T^0\right\} \quad \text{and} \quad \left\{\underline{S}_0^1, \ldots, \underline{S}_T^1\right\},$$

both defined on the filtered belief space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t=0}^T, Bel)$ introduced in Section 5.1. In particular, $\{\underline{S}_0^1, \ldots, \underline{S}_T^1\}$ is assumed to be a DS-multiplicative binomial process as defined in the previous section, with $\underline{S}_0^1 = s_0 > 0$, $u > d > 0$, whose transition belief functions are given by β_k determined by (5.27), with parameters $b_u, b_d > 0$, $b_u + b_d \leq 1$.

On the other hand, the process $\{\underline{S}_0^0, \ldots, \underline{S}_T^0\}$ is deterministic with $\underline{S}_0^0 = 1$ and, for $t = 1, \ldots, T$,

$$\underline{S}_t^0 = (1+r)\underline{S}_{t-1}^0, \tag{5.39}$$

where $r > -1$ is the *risk-free interest rate* over the single period. We further assume that the bond is frictionless, meaning that its lower price $\underline{S}_t^0$ and its upper price $\overline{S}_t^0$ coincide, for all $t = 0, \ldots, T$, then we denote it by S_t^0.

As usual, given the filtered belief space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t=0}^T, Bel)$, a process $\{X_0, \ldots, X_T\}$ defined on such space is said to be *adapted* if X_t is $\mathcal{F}_t$-measurable, for $t = 0, \ldots, T$.

Using the process $\{S_0^0, \ldots, S_T^0\}$ as *numéraire*, we can define the (lower) *discounted process* $\{\underline{S}_0^*, \ldots, \underline{S}_T^*\}$ setting, for $t = 0, \ldots, T$

$$\underline{S}_t^* = \frac{\underline{S}_t^1}{S_t^0} = \frac{\underline{S}_t^1}{(1+r)^t}, \tag{5.40}$$

which is trivially seen to be adapted.

In analogy with Definition 4.2, we define the martingale property in the multi-period setting.

Definition 5.3 *An adapted process $\{X_0, \ldots, X_T\}$ on the filtered belief space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t=0}^T, Bel)$ is said to be a:*

one-step Choquet martingale *if, for $t = 0, \ldots, T-1$, it holds that*

$$\mathbb{C}[X_{t+1}|\mathcal{F}_t] = X_t.$$

one-step Choquet super[sub]-martingale *if, for $t = 0, \ldots, T-1$, it holds that*

$$\mathbb{C}[X_{t+1}|\mathcal{F}_t] \leq [\geq] X_t.$$

global Choquet martingale *if, for every $0 \leq t \leq T-1$ and $1 \leq k \leq T-t$, it holds that*

$$\mathbb{C}[X_{t+k}|\mathcal{F}_t] = X_t.$$

global Choquet super[sub]-martingale *if, for every $0 \leq t \leq T-1$ and $1 \leq k \leq T-t$, it holds that*

$$\mathbb{C}[X_{t+k}|\mathcal{F}_t] \leq [\geq] X_t.$$

The following theorem is the analog under Dempster-Shafer uncertainty of the classical theorem of change of measure for the probabilistic binomial pricing model (Cerný, 2009; Pliska, 1997). In what follows, analogously to probability theory and to Section 4.3, a belief function $\widehat{Bel} : \mathcal{F} \to [0,1]$ is said to be *equivalent* to the belief function Bel if it satisfies Definition 4.1. In particular, since the reference belief function Bel is strictly positive on $\mathcal{F} \setminus \{\emptyset\}$, an equivalent $\widehat{Bel}$ will satisfy the same property.

Note 21: *The assumption of positivity of the belief function Bel follows a financial interpretation as remarked in Note 10. Moreover, it assures that the proposed conditioning rules, in particular the product (geometric) conditioning rule, can be applied. However, from a mathematical point of view, the assumption we made is not innocuous and it may be overcome by considering conditioning rules for belief (or plausibility) functions, as proposed in an axiomatic way in Petturiti and Vantaggi (2022), that generalize the conditioning rules in Section 1.2.3 and take into account events of null measure.*

In what follows $\widehat{\mathbb{C}}[\cdot|\mathcal{F}_t]$ denotes the conditional Choquet expectation with respect to $\widehat{Bel}$.

Theorem 5.2
The condition $u > 1 + r > d > 0$ is necessary and sufficient to the existence of a belief function $\widehat{Bel} : \mathcal{F} \to [0,1]$ equivalent to Bel such that the (lower) discounted process $\{\underline{S}_0^*, \ldots, \underline{S}_T^*\}$ on the filtered belief space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t=0}^T, \widehat{Bel})$ satisfies the following properties:

(a) it is a DS-multiplicative binomial process with transition belief functions $\{\widehat{\beta}_k : k = 1, \ldots, T\}$ satisfying (5.27) with parameters

$$u^* = \frac{u}{1+r},\ d^* = \frac{d}{1+r},\ \widehat{b_u} = \frac{(1+r)-d}{u-d} \text{ and } \widehat{b_d} \in (0, 1 - \widehat{b_u}],$$

(b) it is a *one-step* Choquet martingale, i.e., for $t = 0, \ldots, T-1$ it holds that

$$\widehat{\mathbb{C}}[\underline{S}_{t+1}^*|\mathcal{F}_t] = \underline{S}_t^*,$$

(c) it is a *global* Choquet super-martingale, i.e., for every $0 \le t \le T-1$ and $1 \le k \le T-t$, it holds that

$$\widehat{\mathbb{C}}[\underline{S}_{t+k}^*|\mathcal{F}_t] \le \underline{S}_t^*.$$

Proof. We prove only sufficiency as necessity is readily verified. Hence, suppose $u > 1 + r > d > 0$. Property *(a)* follows immediately, by taking the discounted "up" and "down" coefficients $u^* = \frac{u}{1+r}$ and $d^* = \frac{d}{1+r}$ and taking

$$\widehat{b_u} = \frac{(1+r)-d}{u-d} \quad \text{and} \quad \widehat{b_d} \in (0, 1 - \widehat{b_u}].$$

Property *(b)* follows by Proposition 5.5, noticing that $\{\underline{S}_0^* = s_0^*, \ldots, \underline{S}_t^* = s_t^*\} = \{\underline{S}_0^1 = s_0, \ldots, \underline{S}_t^1 = s_t\}$, since

$$\begin{aligned}
\widehat{\mathbb{C}}[\underline{S}_{t+1}^*|\underline{S}_0^* = s_0^*, \ldots, \underline{S}_t^* = s_t^*] &= d^* s_t^*[\widehat{b_d} + 1 - (\widehat{b_u} + \widehat{b_d})] + u^* s_t^* \widehat{b_u} \\
&= s_t^* \left[\frac{d(u-(1+r))}{(1+r)(u-d)} + \frac{u((1+r)-d)}{(1+r)(u-d)}\right] \\
&= s_t^*.
\end{aligned}$$

We prove property *(c)* by conditioning on $\{\underline{S}_0^* = s_0^*, \ldots, \underline{S}_t^* = s_t^*\}$. By Proposition 5.5, we have that

$$\widehat{\mathbb{C}}[\underline{S}_{t+k}^*|\underline{S}_0^* = s_0^*, \ldots, \underline{S}_t^* = s_t^*] = \sum_{z=0}^{k} \delta_z u^{*z} d^{*k-z} s_t^*,$$

where $\delta_0, \ldots, \delta_k \geq 0$ and $\sum_{z=0}^{k} \delta_z = 1$, and the δ_z's are defined, for $z = 0, \ldots, k$, as

$$\delta_z = \binom{k}{z} \widehat{b_u}^z \widehat{b_d}^{k-z} + \sum_{h=1}^{k-z} \binom{k-h}{z} \widehat{b_u}^z \widehat{b_d}^{k-h-z} (1 - (\widehat{b_u} + \widehat{b_d})),$$

in which the second summation is 0 for $z = k$. Moreover, by well-known results on the classical binomial model we have that

$$s_t^* = \sum_{z=0}^{k} \alpha_z u^{*z} d^{*k-z} s_t^*,$$

where $\alpha_0, \ldots, \alpha_k \geq 0$ and $\sum_{z=0}^{k} \alpha_z = 1$, and the α_z's are defined, for $z = 0, \ldots, k$, as

$$\alpha_z = \binom{k}{z} \widehat{b_u}^z (1 - \widehat{b_u})^{k-z}.$$

If $\widehat{b_d} = 1 - \widehat{b_u}$, then $\delta_z = \alpha_z$, for $z = 0, \ldots, k$, and so

$$\widehat{\mathbb{C}}[\underline{S}_{t+k}^* | \underline{S}_0^* = s_0^*, \ldots, \underline{S}_t^* = s_t^*] = s_t^*.$$

Thus, suppose $\widehat{b_d} \in (0, 1 - \widehat{b_u})$. If $k = 1$, then by property *(b)* we still have that $\widehat{\mathbb{C}}[\underline{S}_{t+1}^* | \underline{S}_0^* = s_0^*, \ldots, \underline{S}_t^* = s_t^*] = s_t^*$. Therefore, suppose $k > 1$. In this case, after a straightforward algebraic manipulation we have that, for $z = 0, \ldots, k-1$,

$$\delta_z = \widehat{b_u}^z \left\{ \sum_{h=1}^{k-z} \left[\binom{k-h+1}{z} - \binom{k-h}{z} \right] \widehat{b_d}^{k-h-z+1} + 1 - \widehat{b_u} \sum_{h=1}^{k-z} \binom{k-h}{z} \widehat{b_d}^{k-h-z} \right\}$$

and $\delta_k = \widehat{b_u}^k$. From this, since $\widehat{b_d} < 1 - \widehat{b_u}$, we get that

$$\delta_0 = 1 - \widehat{b_u} \sum_{h=1}^{k} \widehat{b_d}^{k-h} > 1 - \widehat{b_u} \sum_{h=1}^{k} (1 - \widehat{b_u})^{k-h} = \alpha_0,$$

moreover,

$$\begin{aligned}
\delta_0 + \delta_1 &= 1 - \widehat{b_u} \sum_{h=1}^{k} \widehat{b_d}^{k-h} + \widehat{b_u} \left\{ \sum_{h=1}^{k-1} \widehat{b_d}^{k-h} + 1 - \widehat{b_u} \sum_{h=1}^{k-1} (k-h) \widehat{b_d}^{k-h-1} \right\} \\
&= 1 - \widehat{b_u}^2 \sum_{h=1}^{k-1} (k-h) \widehat{b_d}^{k-h-1} \\
&> 1 - \widehat{b_u}^2 \sum_{h=1}^{k-1} (k-h)(1 - \widehat{b_u})^{k-h-1} = \alpha_0 + \alpha_1.
\end{aligned}$$

More generally, for $j = 0, \ldots, k-2$, we have that

$$\sum_{z=0}^{j} \delta_z = 1 - \widehat{b_u}^{j+1} \sum_{h=1}^{k-j} \binom{k-h}{j} \widehat{b_d}^{k-h-j} > 1 - \widehat{b_u}^{j+1} \sum_{h=1}^{k-j} \binom{k-h}{j} (1 - \widehat{b_u})^{k-h-j} = \sum_{z=0}^{j} \alpha_z,$$

while we get that

$$\begin{aligned}\sum_{z=0}^{k-1} \delta_z = & \quad 1 - \widehat{b_u}^k & = \sum_{z=0}^{k-1} \alpha_z, \\ \sum_{z=0}^{k} \delta_z = & \quad\quad 1 & = \sum_{z=0}^{k} \alpha_z.\end{aligned}$$

Hence, we have shown that $\delta_0, \ldots, \delta_k$ and $\alpha_0, \ldots, \alpha_k$ are probability distributions on $\mathcal{A}_k s_t$ such that $\alpha_0, \ldots, \alpha_k$ first-order stochastically dominates $\delta_0, \ldots, \delta_k$. In turn, this implies that

$$\widehat{\mathbb{C}}[\underline{S}^*_{t+k} | \underline{S}^*_0 = s^*_0, \ldots, \underline{S}^*_t = s^*_t] < s^*_t,$$

and this concludes the proof.

□

Note 22: *Due to the time-homogeneity and Markov properties of the process* $\{\underline{S}^*_0, \ldots, \underline{S}^*_T\}$ *and the fact that* $\{\underline{S}^*_t = s^*_t\} = \{\underline{S}^1_t = s_t\}$, *properties (b) and (c) of Theorem 5.2 reduce to*

(b') for $t = 0, \ldots, T-1$ *it holds that*

$$\widehat{\mathbb{C}}[\underline{S}^*_{t+1} | \underline{S}^*_t] = \underline{S}^*_t,$$

(c') for every $0 \leq t \leq T-1$ *and* $1 \leq k \leq T-t$, *it holds that*

$$\widehat{\mathbb{C}}[\underline{S}^*_{t+k} | \underline{S}^*_t] \leq \underline{S}^*_t.$$

We also have that the original process $\{\underline{S}^1_0, \ldots, \underline{S}^1_T\}$ *continues to be a DS-multiplicative binomial process, seen in the new filtered belief space* $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t=0}^T, \widehat{Bel})$.

Following the usual terminology of mathematical finance, the belief function $\widehat{Bel}$ singled out by the choice of $\widehat{b_u}$ and $\widehat{b_d}$ as in Theorem 5.2, will be called an *equivalent one-step Choquet martingale belief function* or, simply, *one-step risk-neutral belief function.* By contrast, the original belief function Bel will be called *real-world belief function.* We remark that there are actually infinitely many one-step risk-neutral belief functions, depending on the choice of $\widehat{b_d} \in (0, 1 - \widehat{b_u}]$. The adjective risk-neutral for such a belief function $\widehat{Bel}$ is justified by the fact that the Choquet expectation at time t of the return of the stock over the period $[t, t+1]$ coincides with the risk-free return $1 + r$, that is

$$\widehat{\mathbb{C}} \left[\left. \frac{\underline{S}^1_{t+1}}{\underline{S}^1_t} \right| \mathcal{F}_t \right] = 1 + r. \tag{5.41}$$

Note 23: *Asking that the (lower) discounted process* $\{\underline{S}^*_0, \ldots, \underline{S}^*_T\}$ *satisfies one-step Choquet martingale and two-step Choquet martingale, i.e., it holds that*

$$\begin{aligned}\widehat{\mathbb{C}}\left[\underline{S}^*_{t+1} | \mathcal{F}_t\right] &= \underline{S}^*_t, \\ \widehat{\mathbb{C}}\left[\underline{S}^*_{t+2} | \mathcal{F}_t\right] &= \underline{S}^*_t,\end{aligned}$$

then $\widehat{Bel}$ *reduces to a probability measure since* $\widehat{b_u} = \frac{(1+r)-d}{u-d}$ *and* $\widehat{b_d} = 1 - \widehat{b_u}$. *We stress that it coincides with the equivalent martingale measure* Q *of the classical binomial pricing model (see Section 2.2.1).*

The following corollary proves what is stated in Note 23 for every $1 \leq k \leq T-t$ and it is an immediate consequence of the proof of Theorem 5.2.

Corollary 5.1 *If $T > 1$ and $u > 1+r > d > 0$, then the (lower) discounted process $\{\underline{S}_0^*, \ldots, \underline{S}_T^*\}$ satisfying the properties (a)–(c) of Theorem 5.2 further satisfies the property:*

(d) it is a global Choquet martingale, i.e., for every $0 \leq t \leq T-1$ and $1 \leq k \leq T-t$, it holds that

$$\widehat{\mathbb{C}}[\underline{S}_{t+k}^*|\mathcal{F}_t] = \underline{S}_t^*,$$

if and only if $\widehat{b}_d = 1 - \widehat{b}_u$, that is $\widehat{Bel}$ is a probability measure.

5.3 A dynamic pricing rule with bid-ask spreads

Consider the market introduced in Section 5.2, described by the price processes

$$\{S_0^0, \ldots, S_T^0\} \quad \text{and} \quad \{\underline{S}_0^1, \ldots, \underline{S}_T^1\},$$

defined on the real-world filtered belief space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t=0}^T, Bel)$. We recall that $\{\underline{S}_0^1, \ldots, \underline{S}_T^1\}$ is assumed to be a DS-multiplicative binomial process with transition belief function β_k (5.27), while $\{S_0^0, \ldots, S_T^0\}$ is a deterministic process.

We face the problem of finding the lower price of a simple European-type *derivative contract* with maturity T, whose underlying asset is the stock. Such a contract has payoff at the maturity T given by

$$\underline{Y}_T = \varphi(\underline{S}_T^1), \tag{5.42}$$

where φ is a suitable *contract function* defined on the range of $\underline{S}_T^1$. From a financial point of view, the process $\{\underline{Y}_0, \ldots, \underline{Y}_T\}$ can be interpreted as the *lower* price evolution of the derivative with payoff $\underline{Y}_T = \varphi(\underline{S}_T^1)$.

Note 24: *The payoff of the derivative is assumed to be frictionless, i.e., if $\{\overline{Y}_0, \ldots, \overline{Y}_T\}$ is the upper price process of the same derivative, we assume that $\overline{Y}_T = \varphi(\underline{S}_T^1) = \underline{Y}_T = Y_T$.*

Remark 9. We continue to assume that $\widehat{b}_u$ and $\widehat{b}_d$ are defined as in Theorem 5.2, determining the one-step risk-neutral belief function $\widehat{Bel}$ and the corresponding risk-neutral filtered belief space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t=0}^T, \widehat{Bel})$.

We define a lower price process for the derivative contract by setting, for $t = 0, \ldots, T-1$,

$$\underline{Y}_t = \frac{1}{1+r}\widehat{\mathbb{C}}[\underline{Y}_{t+1}|\mathcal{F}_t], \tag{5.43}$$

where $\widehat{\mathbb{C}}[\cdot|\mathcal{F}_t]$ denotes the conditional Choquet expectation with respect to $\widehat{Bel}$. We actually have that, since $Y_T = \varphi(\underline{S}_T^1)$, then $\underline{Y}_t = \varphi_t(\underline{S}_t^1)$ where φ_t is a function on the range of $\underline{S}_t^1$, for $t = 0, \ldots, T-1$, and $\varphi_T = \varphi$, that is all random variables $\underline{Y}_t$'s turn out to be functions of the corresponding random variables $\underline{S}_t^1$'s. In particular, by the time-homogeneity and Markov properties of the process $\{\underline{S}_0^1, \ldots, \underline{S}_T^1\}$ under the one-step risk-neutral belief function $\widehat{Bel}$, we get that

$$\underline{Y}_t = \frac{1}{1+r}\widehat{\mathbb{C}}[\underline{Y}_{t+1}|\underline{S}_t^1]. \tag{5.44}$$

The above construction defines a process $\{\underline{Y}_0, \ldots, \underline{Y}_T\}$, still adapted to the risk-neutral filtered belief space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t=0}^T, \widehat{Bel})$.

Using again the process $\{S_0^0, \ldots, S_T^0\}$ as numéraire, we can define the (lower) discounted process $\{\underline{Y}_0^*, \ldots, \underline{Y}_T^*\}$ setting, for $t = 0, \ldots, T$

$$\underline{Y}_t^* = \frac{\underline{Y}_t}{S_t^0} = \frac{\underline{Y}_t}{(1+r)^t}, \tag{5.45}$$

which is trivially seen to be adapted.

Theorem 5.3
The discounted process $\{\underline{Y}_0^*, \ldots, \underline{Y}_T^*\}$ on the risk-neutral filtered belief space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t=0}^T, \widehat{Bel})$ satisfies the properties:

(a) it is a *one-step* Choquet martingale, i.e., for $t = 0, \ldots, T-1$, it holds that

$$\widehat{\mathbb{C}}[\underline{Y}_{t+1}^* | \mathcal{F}_t] = \underline{Y}_t^*,$$

(b) it is a *global* Choquet super-martingale, i.e., for every $0 \leq t \leq T-1$ and $1 \leq k \leq T-t$, it holds that

$$\widehat{\mathbb{C}}[\underline{Y}_{t+k}^* | \mathcal{F}_t] \leq \underline{Y}_t^*,$$

(c) it is a *global* Choquet martingale, i.e., for every $0 \leq t \leq T-1$ and $1 \leq k \leq T-t$, it holds that

$$\widehat{\mathbb{C}}[\underline{Y}_{t+k}^* | \mathcal{F}_t] = \underline{Y}_t^*,$$

when $\widehat{b_d} = 1 - \widehat{b_u}$.

Proof. Property *(a)* is an immediate consequence of (5.45) and the positive homogeneity property of the conditional Choquet expectation, indeed

$$\begin{aligned}
\underline{Y}_t^* = \frac{\underline{Y}_t}{(1+r)^t} &= \frac{1}{(1+r)^t} \frac{1}{1+r} \widehat{\mathbb{C}}[\underline{Y}_{t+1} | \mathcal{F}_t] \\
&= \widehat{\mathbb{C}}\left[\left. \frac{\underline{Y}_{t+1}}{(1+r)^{t+1}} \right| \mathcal{F}_t \right] = \widehat{\mathbb{C}}[\underline{Y}_{t+1}^* | \mathcal{F}_t].
\end{aligned}$$

We prove property *(b)*. Due to its definition, the (lower) discounted process $\{\underline{Y}_0^*, \ldots, \underline{Y}_T^*\}$ can be expressed as $\underline{Y}_t^* = \psi_t(\underline{S}_t^1)$ for a suitable $\psi_t : \mathcal{S}_t \to \mathbb{R}$, for $t = 0, \ldots, T$, where $\psi_T(\underline{S}_T^1) = \frac{\varphi(\underline{S}_T^1)}{(1+r)^T}$. Fix $0 \leq t \leq T-1$, $1 \leq k \leq T-t$, and $s_t \in \mathcal{S}_t$. By Proposition 5.5, it holds that

$$\begin{aligned}
\widehat{\mathbb{C}}[\underline{Y}_{t+k}^* | \underline{S}_t^1 = s_t] &= \sum_{z=0}^{k} \mu_k'(\{a_z\}) \psi_{t+k}(a_z s_t) \\
&\quad + \sum_{h=1}^{k} \sum_{z=0}^{k-h} \mu_k'([a_z, a_{z+h}]) \min_{a_i \in [a_z, a_{z+h}]} \psi_{t+k}(a_i s_t)
\end{aligned}$$

We also have that, for $h = 0, \ldots, k-1$ and $s_{t+h} \in \mathcal{S}_{t+h}$

$$\begin{aligned}
\psi_{t+h}(s_{t+h}) &= \widehat{b_d} \psi_{t+h+1}(d s_{t+h}) + \widehat{b_u} \psi_{t+h+1}(u s_{t+h}) \\
&\quad + \min\{\psi_{t+h+1}(d s_{t+h}), \psi_{t+h+1}(u s_{t+h})\}(1 - (\widehat{b_u} + \widehat{b_d})).
\end{aligned}$$

Since
$$\min\{\psi_{t+h+1}(ds_{t+h}), \psi_{t+h+1}(us_{t+h})\} \geq \min_{a_i \in \mathcal{A}_{k-h}} \psi_{t+k}(a_i s_{t+h}),$$
starting from $\psi_t(s_t)$, an iterative substitution and minorization shows that
$$\begin{aligned}\psi_t(s_t) &\geq \sum_{z=0}^{k} \mu'_k(\{a_z\})\psi_{t+k}(a_z s_t) \\ &\quad + \sum_{h=1}^{k} \sum_{z=0}^{k-h} \mu'_k([a_z, a_{z+h}]) \min_{a_i \in [a_z, a_{z+h}]} \psi_{t+k}(a_i s_t) \\ &= \widehat{\mathbb{C}}[\underline{Y}^*_{t+k} | \underline{S}^1_t = s_t],\end{aligned}$$
thus the claim follows.

Property *(c)* is an immediate consequence of well-known results on the classical binomial model (see Section 2.2.1).

□

We stress that condition $\widehat{b}_d = 1 - \widehat{b}_u$ is *sufficient* for the (lower) discounted process $\{\underline{Y}^*_0, \ldots, \underline{Y}^*_T\}$ to be a global Choquet martingale but it is not necessary. To see this, it is enough to take a constant contract function φ defined on the range of $\underline{S}^1_T$, for which $\{\underline{Y}^*_0, \ldots, \underline{Y}^*_T\}$ is a global Choquet martingale, independently of $\widehat{b}_d \in (0, 1 - \widehat{b}_u]$.

Since we interpret the undiscounted process $\{\underline{Y}_0, \ldots, \underline{Y}_T\}$ as the lower price evolution of the derivative, such process can be associated with an *upper* price process $\{\overline{Y}_0, \ldots, \overline{Y}_T\}$ under the assumption in Note 24, by setting for $t = 0, \ldots, T-1$,
$$\overline{Y}_t = -\frac{1}{1+r}\widehat{\mathbb{C}}[-\overline{Y}_{t+1} | \mathcal{F}_t]. \tag{5.46}$$
The pair of processes $\{\underline{Y}_0, \ldots, \underline{Y}_T\}$ and $\{\overline{Y}_0, \ldots, \overline{Y}_T\}$ can thus be used to model the time evolution of bid-ask spreads in a market with frictions.

Proposition 5.6
The following statements hold:

(i) $\underline{Y}_t \leq \overline{Y}_t$, for $t = 0, \ldots, T$;

(ii) if φ is *non-decreasing* then the lower price process $\{\underline{Y}_0, \ldots, \underline{Y}_T\}$ does not depend on the choice of $\widehat{b}_d \in (0, 1 - \widehat{b}_u]$;

(iii) if φ is *non-increasing* then the upper price process $\{\overline{Y}_0, \ldots, \overline{Y}_T\}$ does not depend on the choice of $\widehat{b}_d \in (0, 1 - \widehat{b}_u]$.

Proof. Statement *(i)* is an immediate consequence of (5.43) and (5.46). If φ is non-decreasing, statement *(ii)* follows by (5.43) and Proposition 5.5 since, for $t = 0, \ldots, T$, it is easy to show that $\underline{Y}_t = \varphi_t(\underline{S}^1_t)$, where $\varphi_t : \mathcal{S}_t \to \mathbb{R}$ is non-decreasing. If φ is non-increasing, statement *(iii)* follows by (5.46) and Proposition 5.5 since, for $t = 0, \ldots, T$, it is easy to show that $\overline{Y}_t = \varphi_t(\underline{S}^1_t)$, where $\varphi_t : \mathcal{S}_t \to \mathbb{R}$ is non-increasing.
□

Example 5.4 *Let* $T = 3$, $r = 0.04$, $\underline{S}_0^1 = €100$, $u = 1.2$, $d = 0.8$, *and consider a European put option on the stock with maturity* T *and strike price* $\tilde{K} = €100$, *whose final payoff is*

$$P_T = \max\{\tilde{K} - \underline{S}_T^1, 0\}.$$

This is the binomial tree representation of the evolution of the underlying asset:

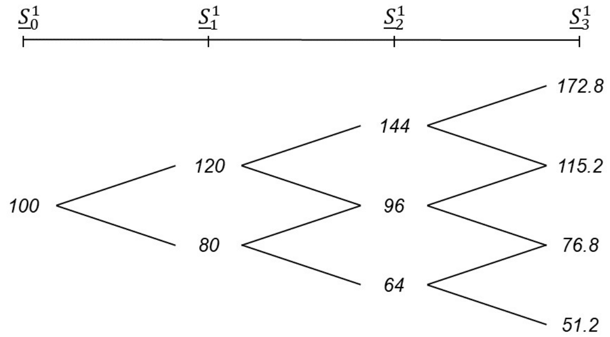

Figure 5.2. (Lower) price process of the underlying asset $\{\underline{S}_0^1, \underline{S}_1^1, \underline{S}_2^1, \underline{S}_3^1\}$.

In this case we have $\widehat{b_u} = 0.6$ *and* $\widehat{b_d} \in (0, 0.4]$.

Figures 5.3–5.4 show, respectively, the lower and the upper price processes of the put option for $\widehat{b_d} = 0.4 \cdot 0.999$.

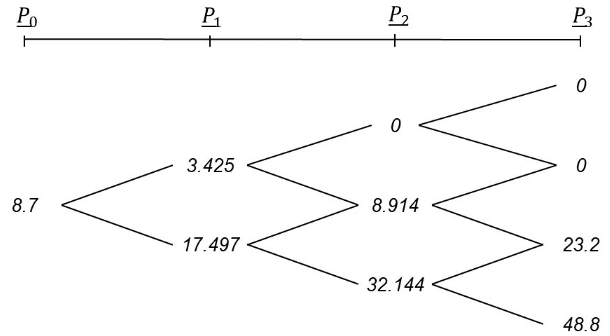

Figure 5.3. Lower price process of the put option $\{\underline{P}_0, \underline{P}_1, \underline{P}_2, \underline{P}_3\}$.

Setting $\widehat{b_d} = 0.4\epsilon$, *we have that*

$$\underline{P}_0 = \frac{11.136\epsilon^2 - 1.3312\epsilon^3}{(1.04)^3} \quad \textit{and} \quad \overline{P}_0 = \frac{3 \cdot 0.6 \cdot 0.4^2 \cdot 23.2 + 0.4^3 \cdot 48.8}{(1.04)^3},$$

where $\overline{P}_0$ *does not depend on* ϵ *by Proposition 5.6.*

Figure 5.5 shows the graph of the bid-ask spread $\overline{P}_0 - \underline{P}_0$ *as a function of* $\epsilon \in (0, 1]$.

♦

We point out that another possibility for defining a lower price process is to set

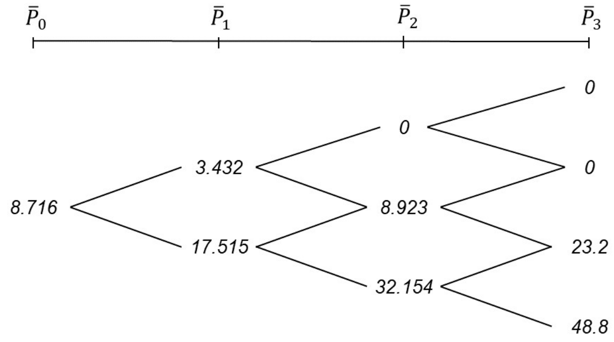

Figure 5.4. Upper price process of the put option $\{\overline{P}_0, \overline{P}_1, \overline{P}_2, \overline{P}_3\}$.

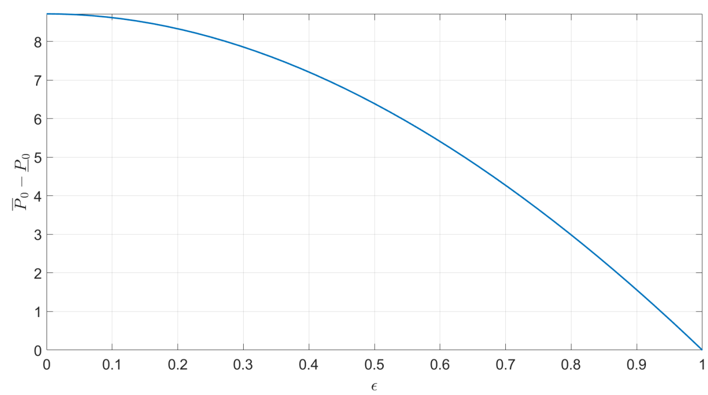

Figure 5.5. Bid-ask spread $\overline{P}_0 - \underline{P}_0$ as a function of $\epsilon \in (0, 1]$.

$\underline{\underline{Y}}_T = \varphi(\underline{S}^1_T)$ and, for $t = 0, \ldots, T-1$, define a *(multi-step) lower price process*

$$\underline{\underline{Y}}_t = \frac{1}{(1+r)^{T-t}} \widehat{\mathbb{C}}\left[\underline{\underline{Y}}_T | \mathcal{F}_t\right]. \tag{5.47}$$

The resulting lower price process $\{\underline{\underline{Y}}_0, \ldots, \underline{\underline{Y}}_T\}$ coincides with $\{\underline{Y}_0, \ldots, \underline{Y}_T\}$ if $\widehat{b_d} = 1 - \widehat{b_u}$, while in general we have $\underline{\underline{Y}}_t \leq \underline{Y}_t$, by virtue of Theorem 5.2. The fact that $\{\underline{\underline{Y}}_0, \ldots, \underline{\underline{Y}}_T\}$ gives rise to a greater dilation (as shown in the following Example 5.5) in lower prices makes us favor the one-step approach given by (5.43).

Example 5.5 *Consider the put option of Example 5.4. Let us compute*

$$\underline{\underline{P}}_0 = \frac{1}{(1+r)^{-3}} \widehat{\mathbb{C}}[P_3 | \mathcal{F}_0]$$

with $\widehat{b_u} = 0.6$ *and* $\widehat{b_d} = 0.4 \cdot 0.999$ *through* (5.47)*. We have that*

$$\underline{\underline{P}}_0 = 8.6976 < \underline{P}_0 = 8.7.$$

Then condition (b) of Theorem 5.3 is satisfied.

♦

Example 5.6 *Let us consider T, r and the lower price process of $\underline{S}^1$ as in Example 5.4, and a straddle option on the stock with maturity T and strike price $\tilde{K} =$ €100, whose final payoff is*

$$Y_T = P_T + C_T = \max\{\tilde{K} - \underline{S}^1_T, 0\} + \max\{\underline{S}^1_T - \tilde{K}, 0\}.$$

As in Example 5.4, we have $\widehat{b_u} = 0.6$ *and* $\widehat{b_d} \in (0, 0.4]$.

Since φ is a non-monotone function, both the lower price process $\{\underline{Y}_0, \underline{Y}_1, \underline{Y}_2, \underline{Y}_3\}$ *and the upper price process* $\{\overline{Y}_0, \overline{Y}_1, \overline{Y}_2, \overline{Y}_3\}$ *depend on the choice of* $\widehat{b_d}$.

Figures 5.6–5.7 show, respectively, the binomial tree representation of the lower and upper price processes for $\widehat{b_d} = 0.4 \cdot 0.999$.

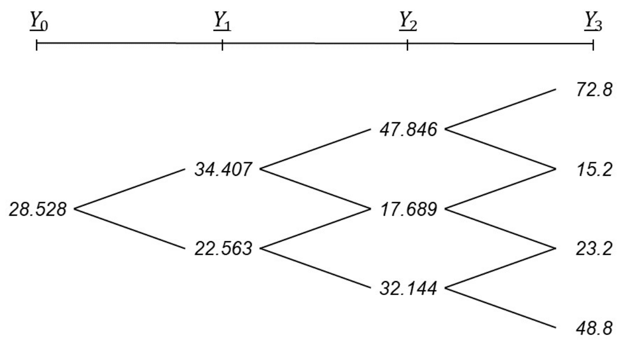

Figure 5.6. Lower price process of the straddle option $\{\underline{Y}_0, \underline{Y}_1, \underline{Y}_2, \underline{Y}_3\}$.

Setting $\widehat{b_d} = 0.4\epsilon$, *we have that*

$$\underline{Y}_0 = \frac{27.6416 + 3.3280\epsilon + 1.1264\epsilon^2}{(1.04)^3},$$

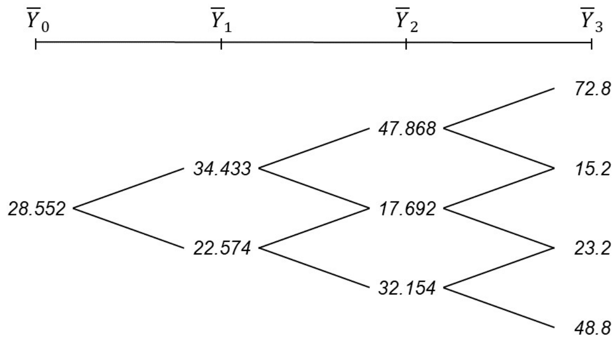

Figure 5.7. Upper price process of the straddle option $\{\overline{Y}_0, \overline{Y}_1, \overline{Y}_2, \overline{Y}_3\}$.

$$\overline{Y}_0 = \frac{24.416 - 6.016(1 - 0.4\epsilon) - 3.2(1 - 0.4\epsilon)^2 + 57.6(1 - 0.4\epsilon)^3}{(1.04)^3}.$$

Figure 5.8 shows the graph of the bid-ask spread $\overline{Y}_0 - \underline{Y}_0$ *as a function of* $\epsilon \in (0, 1]$.

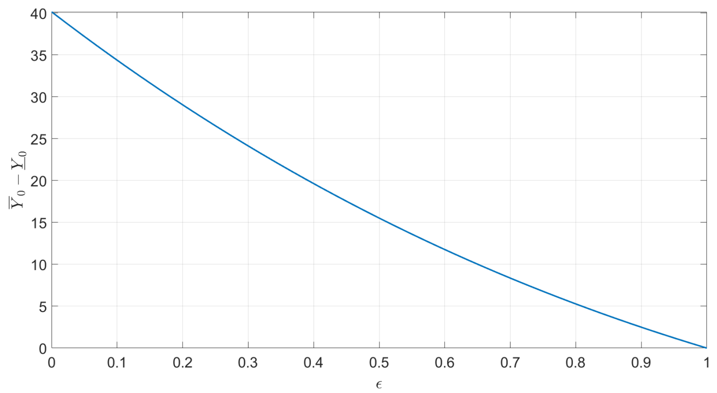

Figure 5.8. Bid-ask spread $\overline{Y}_0 - \underline{Y}_0$ as a function of $\epsilon \in (0, 1]$.

Consider the lower price $\underline{\underline{Y}}_0 = \frac{1}{(1+r)^3}\widehat{\mathbb{C}}[\underline{\underline{Y}}_3|\mathcal{F}_0]$ *(we stress that* $\underline{\underline{Y}}_3 = \underline{Y}_3 = Y_3$*). For* $\widehat{b_d} = 0.4 \cdot 0.999$*, we have that*

$$\underline{\underline{Y}}_0 = 28.5239 < \underline{Y}_0 = 28.5283.$$

In analogy, the upper price $\overline{\overline{Y}}_0 = \frac{1}{(1+r)^3} - \widehat{\mathbb{C}}[-\overline{\overline{Y}}_3]$*, where* $\overline{\overline{Y}}_3 = Y_3$*, is*

$$\overline{\overline{Y}}_0 = 28.5716 > \overline{Y}_0 = 28.5519.$$

Hence, condition (b) of Theorem 5.3 is satisfied.

♦

On the other hand, the (multi-step) lower price process defined through (5.47) assures that a dynamic version of the *put-call parity relation* introduced in Cerreia-Vioglio et al. (2015) (see Section 3.1.2) is satisfied. Indeed, denoting by

$C_T = \max\{\underline{S}_T^1 - \tilde{K}, 0\}$ and $P_T = \max\{\tilde{K} - \underline{S}_T^1, 0\}$ the payoffs of European call and put options on $\underline{S}_T^1$ with strike price $\tilde{K}$, the decomposition

$$C_T - P_T = \underline{S}_T^1 - \tilde{K} \tag{5.48}$$

and the comonotonic additivity of $\widehat{\mathbb{C}}[\cdot|\mathcal{F}_t]$ imply that

$$\widehat{\mathbb{C}}[C_T|\mathcal{F}_t] + \widehat{\mathbb{C}}[-P_T|\mathcal{F}_t] = \widehat{\mathbb{C}}[\underline{S}_T^1|\mathcal{F}_t] - \tilde{K} \tag{5.49}$$

which, after discounting, reduces to

$$\underline{\underline{C}}_t + \frac{\widehat{\mathbb{C}}[-P_T|\mathcal{F}_t]}{(1+r)^{T-t}} = \underline{\underline{S}}_t^1 - \frac{\tilde{K}}{(1+r)^{T-t}}, \tag{5.50}$$

where $\underline{\underline{C}}_t, \underline{\underline{S}}_t^1$ refer to (5.47). Let us stress that, since the lower stock price process under $\widehat{Bel}$ is only a global Choquet super-martingale and a one-step Choquet martingale, we actually have that $\underline{\underline{S}}_t^1 \leq \underline{S}_t^1$.

However, under ambiguity, different forms of put-call parity relations arise: for instance, the form introduced in Chateauneuf et al. (1996) (see (CPP) p. 54) is generally not satisfied in our framework, as it holds only if it is a Choquet-Šipoš integral, which does not allow the presence of frictions, as pointed out in Section 3.1.2.

Note 25: *The proposed DS-multiplicative binomial process could be analogously defined in terms of the wider framework of* 2*-monotone capacities, instead of belief functions, since properties* (1.21), (1.27) *continue to hold. However, departing from the additive probability setting, the tower property and the Chapman-Kolmogorov equation do not hold and, also working with* 2*-monotone capacities, it is required to characterize the whole family of k-step transition* 2*-monotone capacities. Moreover, the one-step time-homogeneity in* (5.16)–(5.17) *should be defined in terms of* 2*-monotone capacities as well as the reduction to the probability model as a special case. It means that the generalized theorem of change of measure (Theorem 5.2) would lead to the same characterization of b_u and b_d in order to have (in the hypothetical new framework) a one-step risk-neutral* 2*-monotone capacity. The multi-step lower pricing rule (that we do not take into account because of the wider dilation of prices) would be different, but the one-step lower pricing rule would remain unchanged. It results that the wider framework of* 2*-monotone capacities would not have significant benefits but certainly more computational difficulties.*

5.4 A dynamic generalized no-arbitrage principle

The construction carried out in the previous section subsumes the classical linear formulation, obtained when we restrict to work with additive belief functions. In this case, we get back to probability theory where the conditional Choquet expectation operator defined in (5.37) reduces to the classical conditional expectation operator, which is linear and satisfies the tower property.

Remark 10. We recall that the classical binomial pricing model builds upon the assumption of a perfect (frictionless and competitive) market under the classical no-arbitrage principle. In particular, the hypothesis of absence of frictions in the market implies that the lower and upper price processes $\{\underline{S}_0^1, \dots, \underline{S}_T^1\}$ and $\{\overline{S}_0^1, \dots, \overline{S}_T^1\}$ always coincide. Such processes, that are considered to be defined on the real-world filtered probability space, can be seen as a special case of a filtered additive belief space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t=0}^T, Bel)$ where $b_u \in (0,1)$ and $b_d = 1 - b_u$, i.e., Bel reduces to a strictly positive probability measure P.

We continue to denote a *portfolio* or *trading strategy* as a bivariate stochastic process $\{\boldsymbol{\lambda}_0, \dots, \boldsymbol{\lambda}_{T-1}\}$ with $\boldsymbol{\lambda}_t = (\lambda_t^0, \lambda_t^1)$, and with $\{V_0, \dots, V_T\}$ the corresponding *portfolio price process* (see (2.13)–(2.14), p.33, with $K = 1$).

The additive formulation of the market above is *dynamically complete* as shown in Proposition 2.1, in particular, for $K = 1$, every simple European-type derivative with payoff $Y_T = \varphi(S_T^1)$ can be replicated by a self-financing strategy $\{\boldsymbol{\lambda}_0, \dots, \boldsymbol{\lambda}_{T-1}\}$. Also in this case, the process $\{Y_0, \dots, Y_T\}$ is interpreted as the price evolution of the derivative, as we work in a frictionless market.

Note 26: *In the classical binomial model of Section 2.2.1, we can interpret $Q \sim P$ as an additive belief function $\widehat{Bel}$ with $\widehat{b_u} = \frac{(1+r)-d}{u-d}$ and $\widehat{b_d} = 1 - \widehat{b_u}$. In this case, the risk-neutral conditional one-step Choquet expectation operator $\widehat{\mathbb{C}}[\cdot|\mathcal{F}_t]$ reduces to a classical one-step expectation operator $\widehat{\mathbb{E}}[\cdot|\mathcal{F}_t]$ and we have that, for $t = 0, \dots, T-1$*

$$Y_t = \frac{1}{1+r}\widehat{\mathbb{E}}[Y_{t+1}|\mathcal{F}_t]. \tag{5.51}$$

Additionally, in the classical setting, the discounted process $\{Y_0^*, \dots, Y_T^*\}$ is a global (Choquet) martingale and, so, we have that (2.28) holds. The classical construction recalled above is intrinsically based on the additivity of Bel and $\widehat{Bel}$, that in Section 2.2.1 are denoted, respectively, as P and Q. Indeed, in case of additive belief functions, the one-step Markov and time-homogeneity properties imply the general Markov and time-homogeneity properties. The same holds for the one-step martingale and global martingale properties.

Since real markets are quite far from being perfect as they can show frictions, mainly in the form of bid-ask spreads (see Section 3.1.2), allowing frictions in the market, i.e., giving up on the additivity of Bel and $\widehat{Bel}$, the above construction necessarily breaks down since $\widehat{\mathbb{C}}[\cdot|\mathcal{F}_t]$ is not linear and does not satisfy the tower property. In financial terms, the lack of linearity of $\widehat{\mathbb{C}}[\cdot|\mathcal{F}_t]$ translates in the lack of duality between the direct definition of the lower price process $\{\underline{Y}_0, \dots, \underline{Y}_t\}$ as a global discounted Choquet expectation and the replicating portfolio representation.

Remark 11. The failure of the tower property in the belief framework implies that working on single periods $[t, t+1]$ is not equivalent to working on larger periods as it is when the additivity is satisfied. Then, the (multi-step) "classical" pricing rule in (2.28) cannot be applied when we work in terms of (non-additive) belief function. In fact, the (one-step) lower pricing rule (5.43) is not equivalent to the (multi-step) lower pricing rule (5.47).

We provide a detailed analysis of implications due to the lack of additivity. If we assume Bel and $\widehat{Bel}$ are non-additive belief functions, i.e., $b_d \in (0, 1 - b_u)$ and $\widehat{b_d} \in (0, 1 - \widehat{b_u})$, then we can still define the lower price process $\{\underline{Y}_0, \dots, \underline{Y}_T\}$ of a simple derivative with payoff $Y_T = \varphi(S_T^1)$ through (5.43), for which we have that

$\underline{Y}_t = \varphi_t(\underline{S}_t^1)$ with $\varphi_t : \mathcal{S}_t \to \mathbb{R}$. In order to have a replicating strategy, in every period $[t, t+1]$, working conditionally on the history of the stock lower price process up to time t, the random vector $\boldsymbol{\lambda}_t = (\lambda_t^0, \lambda_t^1)$ must be chosen by solving the linear system

$$\begin{cases} \lambda_t^0 \underline{S}_{t+1}^0 + \lambda_t^1 u \underline{S}_t^1 = \varphi_{t+1}(u\underline{S}_t^1), \\ \lambda_t^0 \underline{S}_{t+1}^0 + \lambda_t^1 d \underline{S}_t^1 = \varphi_{t+1}(d\underline{S}_t^1), \end{cases} \tag{5.52}$$

where we recall that $\underline{S}_t^0 = \overline{S}_t^0 = S_t^0 = (1+r)^t$. The linear system has a unique solution. In turn, the replication constraint can be compactly rewritten as

$$\lambda_t^0 \underline{S}_{t+1}^0 + \lambda_t^1 \underline{S}_{t+1}^1 = \underline{Y}_{t+1}. \tag{5.53}$$

The lack of linearity of $\widehat{\mathbb{C}}[\cdot|\mathcal{F}_t]$ implies that the resulting trading strategy $\{\boldsymbol{\lambda}_0, \ldots, \boldsymbol{\lambda}_T\}$ is generally not self-financing. Denoting the (lower) price process as $\{\underline{V}_0, \ldots, \underline{V}_T\}$, where $\underline{V}_t = \lambda_t^0 \underline{S}_t^0 + \lambda_t^1 \underline{S}_t^1$, we may have

$$\begin{aligned} \underline{Y}_t &= \frac{1}{1+r}\widehat{\mathbb{C}}[\underline{Y}_{t+1}|\mathcal{F}_t] \\ &= \frac{1}{1+r}\widehat{\mathbb{C}}[\lambda_t^0 \underline{S}_{t+1}^0 + \lambda_t^1 \underline{S}_{t+1}^1|\mathcal{F}_t] \\ &= \frac{1}{1+r}\left(\lambda_t^0(1+r)\underline{S}_t^0 + \widehat{\mathbb{C}}[\lambda_t^1 \underline{S}_{t+1}^1|\mathcal{F}_t]\right) \\ &= \frac{1}{1+r}\left(\lambda_t^0(1+r)^{t+1} + \widehat{\mathbb{C}}[\lambda_t^1 \underline{S}_{t+1}^1|\mathcal{F}_t]\right) \\ &= \lambda_t^0(1+r)^t + \frac{1}{1+r}\widehat{\mathbb{C}}[\lambda_t^1 \underline{S}_{t+1}^1|\mathcal{F}_t] \\ &= \lambda_t^0 \underline{S}_t^0 + \frac{1}{1+r}\widehat{\mathbb{C}}[\lambda_t^1 \underline{S}_{t+1}^1|\mathcal{F}_t] \\ &\neq \lambda_t^0 \underline{S}_t^0 + \lambda_t^1 \underline{S}_t^1 = \underline{V}_t, \end{aligned}$$

where $\widehat{\mathbb{C}}[\lambda_t^1 \underline{S}_{t+1}^1|\mathcal{F}_t] \neq \lambda_t^1 \widehat{\mathbb{C}}[\underline{S}_{t+1}^1|\mathcal{F}_t]$ unless $\lambda_t^1 \geq 0$ (see *(Ch.3)* p.19). This shows that we generally lose the replicating strategy representation of the lower price process.

On the other hand, by virtue of Theorem 5.3, the failure of the tower property of $\widehat{\mathbb{C}}[\cdot|\mathcal{F}_t]$ implies that the (lower) discounted process $\{\underline{Y}_0^*, \ldots, \underline{Y}_T^*\}$ is only a one-step Choquet martingale and a global Choquet super-martingale, but it is generally not a global Choquet martingale. In particular, we only have that

$$\underline{Y}_0 \geq \frac{1}{(1+r)^T}\widehat{\mathbb{C}}[\underline{Y}_T|\mathcal{F}_0]. \tag{5.54}$$

We now investigate further how the choice of $\widehat{b_d} \in (0, 1-\widehat{b_u})$ can be justified from a normative point of view. Indeed, as already highlighted, the classical no-arbitrage principle is inconsistent with this choice, as the only admissible choice is to set $\widehat{b_d} = 1 - \widehat{b_u}$.

To see this, we reformulate the no-arbitrage condition restricting to every single period $[t, t+1]$. At this aim, working conditionally on the history $\{\underline{S}_0^1 = s_0, \ldots, \underline{S}_t^1 = s_t\}$, we can define the events $U(s_t) = \{\underline{S}_{t+1}^1 = us_t\}$ and $D(s_t) = \{\underline{S}_{t+1}^1 = ds_t\}$, which are functions of the value $\underline{S}_t^1$ can take, thus we can write $U(S_t^1)$ and $D(S_t^1)$ to stress this fact. In turn, the one-period market formed by the bond and the stock over $[t, t+1]$ can be augmented by adding the artificial securities whose payoff at time $t+1$ is

$$A_{t+1}^u = \mathbf{1}_{U(\underline{S}_t^1)} \quad \text{and} \quad A_{t+1}^d = \mathbf{1}_{D(\underline{S}_t^1)}, \tag{5.55}$$

that turn out to be *Arrow-Debreu securities* (Cerný, 2009). Pricing through (5.43), the prices at time t of Arrow-Debreu securities are set equal to

$$A_t^u = \frac{1}{1+r}\widehat{\mathbb{C}}[A_{t+1}^u|\mathcal{F}_t] = \frac{\widehat{b_u}}{1+r}, \tag{5.56}$$

$$A_t^d = \frac{1}{1+r}\widehat{\mathbb{C}}[A_{t+1}^d|\mathcal{F}_t] = \frac{\widehat{b_d}}{1+r}. \tag{5.57}$$

In the augmented one-period market over $[t, t+1]$, a portfolio is a vector $\boldsymbol{\delta}_t = (\delta_t^0, \delta_t^1, \delta_t^2, \delta_t^3)$, where the δ_t^i's are $\mathcal{F}_t$-measurable random variables expressing, respectively, the number of units of bond, stock and Arrow-Debreu's securities to buy (if positive) or short-sell (if negative) at time t up to time $t+1$. Furthermore, we can define a local (lower) price process $\{\underline{\Pi}_t, \underline{\Pi}_{t+1}\}$ associated with $\boldsymbol{\delta}_t$ over $[t, t+1]$ by defining the random variables

$$\underline{\Pi}_t = \delta_t^0 \underline{S}_t^0 + \delta_t^1 \underline{S}_t^1 + \delta_t^2 A_t^u + \delta_t^3 A_t^d, \tag{5.58}$$

$$\underline{\Pi}_{t+1} = \delta_t^0 \underline{S}_{t+1}^0 + \delta_t^1 \underline{S}_{t+1}^1 + \delta_t^2 A_{t+1}^u + \delta_t^3 A_{t+1}^d. \tag{5.59}$$

Given the history $\{\underline{S}_0^1 = s_0, \ldots, \underline{S}_t^1 = s_t\}$, $U(s_t)$ and $D(s_t)$ form a partition of each event $\{\underline{S}_t^1 = s_t\}$, moreover, the random variables $\underline{\Pi}_t$ and $\underline{\Pi}_{t+1}$ can be simply regarded as functions with domain $\mathcal{W}(s_t) = \{U(s_t), D(s_t)\}$, where $\underline{\Pi}_t$ is actually constant over $\mathcal{W}(s_t)$.

If we are at time t, the tacit assumption of the classical no-arbitrage condition concerning time $t+1$ is to work under *completely resolving uncertainty.* This means that, given the history $\{\underline{S}_0^1 = s_0, \ldots, \underline{S}_t^1 = s_t\}$, at time $t+1$ the market agent will be always able to determine which one between the mutually exclusive events $U(s_t)$ and $D(s_t)$ has occurred. In this setting, the *one-step arbitrage opportunity* is a portfolio $\boldsymbol{\delta}_t$ that satisfies one of the following two conditions, where comparisons are intended over $\mathcal{W}(s_t)$ given the history $\{\underline{S}_0^1 = s_0, \ldots, \underline{S}_t^1 = s_t\}$:

(i) $\underline{\Pi}_t < 0$ and $\underline{\Pi}_{t+1} = 0$;

(ii) $\underline{\Pi}_t \leq 0$ and $\underline{\Pi}_{t+1} \geq 0$ with $\underline{\Pi}_{t+1} \neq 0$.

It is well known that the absence of one-step arbitrage opportunities is equivalent to $u > 1 + r > d > 0$, $\widehat{b_u} = \frac{(1+r)-d}{u-d}$ and $\widehat{b_d} = 1 - \widehat{b_u}$. In turn, this is equivalent to the existence of a unique strictly positive additive risk-neutral belief function $\widehat{Bel}$ that reduces to the already quoted probability measure Q.

Hence, choosing $\widehat{b_d} \in (0, 1 - \widehat{b_u})$ we can always build a one-step arbitrage opportunity. Therefore, to justify the choice of $\widehat{b_d} \in (0, 1 - \widehat{b_u})$ from a normative point of view, we need to generalized the one-step no-arbitrage condition by working under *partially resolving uncertainty*, as done in Chapter 4. The concept of partially resolving uncertainty (Jaffray, 1989) in the present context means that, given the history $\{\underline{S}_0^1 = s_0, \ldots, \underline{S}_t^1 = s_t\}$, at time $t+1$ the market agent may not be able to determine which one between the two mutually exclusive events $U(s_t)$ and $D(s_t)$ has occurred. Thus, she/he needs to consider the set of all the possible pieces of information she/he may acquire once uncertainty is resolved at time $t+1$ which form the set $\mathcal{U}(s_t) = \{U(s_t), D(s_t), U(s_t) \cup D(s_t)\}$.

To address partially resolving uncertainty, the local (lower) price process needs to be changed to $\{\tilde{\underline{\Pi}}_t, \tilde{\underline{\Pi}}_{t+1}\}$ by defining its components as functions defined over $\mathcal{U}(s_t)$ instead of over $\mathcal{W}(s_t)$, given the history up to time t. To do so, we notice that, given the history up to time t, $\underline{S}_t^0, \underline{S}_t^1, A_t^u, A_t^d$ as well as $\underline{S}_{t+1}^0, \underline{S}_{t+1}^1, A_{t+1}^u, A_{t+1}^d$ can be seen as functions with domain $\mathcal{W}(s_t)$. As usual, given a function X defined on $\mathcal{W}(s_t)$, we assume that the market agent adopts a *systematically pessimistic behaviour* under partially resolving uncertainty, considering in place of X the quantity $[X]^{\mathbf{L}}$, as defined in (4.9), that, in this setting is defined on $\mathcal{U}(s_t)$, where, for every $E \in \mathcal{U}(s_t)$, is

$$[X]^{\mathbf{L}}(E) = \min\{X(F) \,:\, F \subseteq E, F \in \mathcal{W}(s_t)\}. \tag{5.60}$$

We finally define

$$\begin{aligned} \tilde{\underline{\Pi}}_t &= \delta_t^0[S_t^0]^{\mathbf{L}} + \delta_t^1[S_t^1]^{\mathbf{L}} + \delta_t^2[A_t^u]^{\mathbf{L}} + \delta_t^3[A_t^d]^{\mathbf{L}}, && (5.61)\\ \tilde{\underline{\Pi}}_{t+1} &= \delta_t^0[S_{t+1}^0]^{\mathbf{L}} + \delta_t^1[S_{t+1}^1]^{\mathbf{L}} + \delta_t^2[A_{t+1}^u]^{\mathbf{L}} + \delta_t^3[A_{t+1}^d]^{\mathbf{L}}. && (5.62) \end{aligned}$$

Also in this case, $\tilde{\underline{\Pi}}_t$ is actually constant over $\mathcal{U}(s_t)$.

In agreement with Chapter 4, we define a *generalized one-step arbitrage opportunity* a portfolio $\boldsymbol{\delta}_t$ that satisfies one of the following two conditions, where comparisons are intended over $\mathcal{U}(s_t)$ given the history $\{\underline{S}_0^1 = s_0, \ldots, \underline{S}_t^1 = s_t\}$:

(i') $\tilde{\underline{\Pi}}_t < 0$ and $\tilde{\underline{\Pi}}_{t+1} \geq 0$ with $\tilde{\underline{\Pi}}_{t+1} = 0$ over $\mathcal{W}(s_t)$;

(ii') $\tilde{\underline{\Pi}}_t \leq 0$ and $\tilde{\underline{\Pi}}_{t+1} \geq 0$ with $\tilde{\underline{\Pi}}_{t+1} \neq 0$ over $\mathcal{W}(s_t)$.

As an immediate consequence of Theorem 4.5 in Section 4.3, avoiding generalized one-step arbitrage opportunities is equivalent to the existence of a conditional belief function $\widehat{Bel}(\cdot|\underline{S}_0^1 = s_0, \ldots, \underline{S}_t^1 = s_t)$ defined on the ring generated by $\mathcal{W}(s_t)$ such that

$$\begin{aligned} \widehat{Bel}(\underline{S}_{t+1}^1 = us_t|\underline{S}_0^1 = s_0, \ldots, \underline{S}_t^1 = s_t) &= \widehat{b_u}, && (5.63)\\ \widehat{Bel}(\underline{S}_{t+1}^1 = ds_t|\underline{S}_0^1 = s_0, \ldots, \underline{S}_t^1 = s_t) &= \widehat{b_d}, && (5.64)\\ \frac{1}{1+r}\widehat{\mathbb{C}}[\underline{S}_{t+1}^1|\underline{S}_0^1 = s_0, \ldots, \underline{S}_t^1 = s_t] &= s_t. && (5.65) \end{aligned}$$

In other terms, the choice of $\widehat{b_d} \in (0, 1-\widehat{b_u})$ is consistent with the generalized one-step no-arbitrage condition, i.e., it does not produce generalized one-step arbitrage opportunities. It is important to notice that abandoning additivity we lose the self-financing property and, therefore, dynamic completeness. In the additive case the one-step no-arbitrage principle alone assures the uniqueness of the global Q defined on the whole $\mathcal{F}$; on the other hand, this is not the case for the generalized one-step no-arbitrage principle since we generally have infinitely many non-additive risk-neutral belief functions $\widehat{Bel}$ compatible with the fixed one-step transition belief functions.

Main results

- The n-nomial market model is incomplete, i.e., there exists a (non-closed) set of equivalent martingale measures $\mathcal{Q}$, whose closure is characterized by the set of its extreme points;
- The lower envelope $\underline{Q}$ of the closure of the set of equivalent martingale measures is proved to be a belief function;
- The closure of the set of equivalent martingale measures does not generally coincide with the core of the lower envelope (belief function) $\underline{Q}$. It follows that the (discounted) Choquet expectation with respect to $\underline{Q}$ does not coincide with the lower expectation with respect to $\mathrm{cl}(\mathcal{Q})$;
- A generalized no-Dutch book condition and no-arbitrage condition are proposed and a generalized first fundamental theorem of asset pricing is proved;
- Inner approximation procedures of the lower envelope are suggested in order to get a risk-neutral belief function such that its ϵ-contamination with respect to $\mathrm{cl}(\mathcal{Q})$ is an equivalent inner approximating Choquet martingale belief function. This leads to a (lower) pricing rule which is able to embody bid-ask spreads;
- In the multi-period setting, a multiplicative binomial random process is characterized in terms of belief functions (called DS-multiplicative binomial process);
- Since the Chapman-Kolmogorov equation (and then the tower property) does not hold, the characterization of the whole set of transition belief functions β_k is required. The choice of this particular β_k is due to its interpretation as a generalized binomial distribution;
- In a market model composed by one frictional risky asset (whose lower price process is modeled by the proposed DS-multiplicative binomial process), a generalized theorem of change of measure is proved;
- A dynamic lower pricing rule which accounts bid-ask spreads is characterized and it is proved that it is a one-step Choquet martingale and a global Choquet super-martingale;
- The reference belief function is finally showed to be consistent with the absence of generalized arbitrage opportunities in every one-step time interval.

Conclusion

This thesis addresses the issue of frictions, largely shown to exist in real markets. After analysing non-linear pricing rules that are proposed in the literature, in order to characterize lower (upper) prices, specifically by means of the Choquet integral, this thesis investigates a one-period n-nomial market model composed by a riskless bond and a risky asset that does not pay dividends. Since it is an incomplete market, the properties of the class of equivalent martingale measures are investigated, and the lower envelope of the class is shown to be a belief function in Dempster-Shafer theory. This suggests a characterization of a lower pricing rule as a (discounted) Choquet expectation. However, two issues arise: the frictional setting violates the classical no-arbitrage principle and the class of equivalent martingale measures is generally a subset of the core of its lower envelope (belief function).

The one-period pricing problem is reformulated in the belief functions framework in order to generalize the (one-step) no-Dutch book condition and the (one-step) no-arbitrage condition, under the paradigm of partially resolving uncertainty. We first prove a generalized first fundamental theorem of asset pricing and show that the (discounted) Choquet expectation with respect to the lower envelope of the class of equivalent martingale measures does not satisfy the no-Dutch book condition. Then, we propose a lower pricing rule relying on the ϵ-contamination technique.

The ϵ-contamination of a reference equivalent martingale measure with respect to the class of equivalent martingale measures results in a lower envelope that continues to fail the no-Dutch book condition. To address this issue, we consider the ϵ-contamination of a reference equivalent martingale measure with a belief function that is an inner approximation of the lower envelope and such that it satisfies the generalized no-Dutch book condition.

A pricing rule based on the (discounted) Choquet expectation is provided and it is shown to be able to embody bid-ask spreads. Then, we consider a multi-period binomial pricing model composed by a riskless and frictionless bond and a frictional asset whose lower price process uncertainty is modeled by a belief function.

The (lower) price process is characterized through a specific family of k-step transition belief functions which make the process a Markov and time-homogeneous process. The choice of this reference belief function is justified by its mathematical tractability and interpretation as a generalization of the binomial process under (additive) probability measures.

This characterization allows us to introduce a dynamic lower pricing rule by means of the (discounted) conditional Choquet expectation. Thus, a generalized theorem of change of measure is proved, in order to assess an equivalent one-step Choquet martingale belief function. Given the payoff of European type derivative whose underlying asset is the frictional one, a dynamic (lower) pricing rule that accounts bid-ask spreads is proposed and the lower price process of the derivative results to be a one-step Choquet martingale. Since the classical no-arbitrage principle is consistent with our model only if the belief function reduces to a probability measure, we conclude with a normative justification, proving that the proposed reference belief function is consistent with the absence of generalized arbitrage opportunities, in every one-step time interval.

The present model, though simple, can be easily calibrated on market data, due to its significant parameterization. Nevertheless, a natural evolution of the present thesis could aim at looking for more complex models, whose development is reserved to the future. In particular, some of the possible future expansions are:

- characterize the DS-multiplicative binomial process considering other conditioning rules, in particular taking into account events with *zero belief* (Petturiti and Vantaggi, 2022);
- define a more complex model of that proposed in Section 5.2, where *more stocks* evolve as DS-multiplicative binomial processes. This would require to express *dependencies* between the processes by referring, for instance, to a suitable notion of correlation for belief functions (Jiang, 2018);
- characterize a DS-multiplicative n-*nomial process*, i.e., the stock is allowed to have $n > 2$ future developments after one-step. This would require to define and characterize a suitable family of transition belief function, in analogy with (5.27);
- study the convergence to continuous time of the proposed DS-multiplicative binomial process. This requires to adopt a suitable convergence criterion (see, e.g., Feng and Nguyen, 2007) in order to get a sort of DS-geometric Brownian motion mimicking the limit derivation of the model due to Black and Scholes (1973).

Bibliography

Allingham, M. (1991). *Arbitrage.* Palgrave Macmillan.

Amihud, Y. and Mendelson, H. (1986). Asset pricing and the bid-ask spread. *Journal of Financial Economics*, 17: 223–249.

Amihud, Y. and Mendelson, H. (1991). Liquidity, Maturity, and the Yields on U.S. Treasury Securities. *The Journal of Finance*, 46(4): 1411–1425.

Aouani, Z., Chateauneuf, A. and Ventura, C. (2021). Propensity for hedging and ambiguity aversion. *Journal of Mathematical Economics*, 97, 102543.

Banz, R.W. (1981). The relationship between return and market value of common stocks. *Journal of Financial Economics*, 9: 3–18.

Bastianello, L., Chateauneuf, A. and Cornet, B. (2022). Put-Call Parities, absence of arbitrage opportunities and non-linear pricing rules. *arXiv:2203.16292*

Ben-Israel, A. (2001). Motzkin's transposition theorem, and the related theorems of Farkas, Gordan and Stiemke. *Encyclopedia of Mathematics, Supplement III*, Kluwer Academic Publisher, Dordrecht.

Bensaïd, B., Lesne, J-P., Pages, H. and Scheinkman, J. (1992). Derivative asset pricing with transaction costs. *Mathematical Finance*, 2(2): 63–86.

Bertsimas, D., Kogan, L. and Lo, A.W. (2001). Hedging Derivative Securities and Incomplete Markets: An ϵ-Arbitrage Approach. *Operations Research*, 49(3): 372–397.

Billingsley, P. (1995). *Probability and measure* (3rd ed.). New York: Wiley.

Bjork, T. (2009). *Arbitrage Theory in Continuous Time.* (3rd ed.). Oxford University Press.

Black, F. and Scholes, M. (1973). The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, 8: 637–654.

Boivin, C. (2019). *Introduction to Belief Functions: The Monty Hall Game.* 10.13140/RG.2.2.14018.53446

Boyle, P.P. and Vorst, T. (1992). Option Replication in Discrete Time with Transaction Costs. *The Journal of Finance*, 47(1): 271–293.

Carr, P., Geman, H. and Madan, D.B. (2001). Pricing and hedging in incomplete markets. *Journal of Financial Economics*, 62(1): 131–167.

Censor, Y. and Zenios, S.A. (1997). *Parallel Optimization: Theory, Algorithms and Applications.* Oxford University Press.

Cerný, A. (2009). *Mathematical Techniques in Finance: Tools for Incomplete Markets.* Princeton: Princeton University Press.

Cerreia-Vioglio, S., Maccheroni, F. and Marinacci, M. (2015). Put-Call Parity and market frictions. *Journal of Economic Theory*, 157: 730–762.

Chateauneuf, A. (1991). On the use of capacities in modeling uncertainty aversion and risk aversion. *Journal of Mathematical Economics*, 20(4): 343–369.

Chateauneuf, A. and Cornet, B. (2022a). Submodular financial markets with frictions. *Economic Theory. Society for the Advancement of Economic Theory (SAET)*, 73(2): 721–744.

Chateauneuf, A. and Cornet, B. (2022b). The risk neutral non-additive probability with market frictions. *Economic Theory Bulletin*, 10(1): 13–25.

Chateauneuf, A. and Jaffray, J.Y. (1989). Some characterizations of lower probabilities and other monotone capacities through the use of Möbius inversion. *Mathematical Social Sciences*, 17(3): 263–283.

Chateauneuf, A., Kast, R. and Lapied, A. (1996). Choquet pricing for financial markets with frictions. *Mathematical Finance*, 6(3): 323–330.

Choquet, G. (1953). Theory of capacities. *Annales de l'Institut Fourier*, 5: 131–295.

Cinfrignini, A. (2022). Pricing through the Choquet integral. *Annali del dipartimento di Metodi e Modelli per l'Economia, la Finanza e il Territorio*, *On-line first Special Issue PhD-2022*: 1–17, ISSN 2611-6634.

Cinfrignini, A., Petturiti, D. and Vantaggi, B. (2022). Markov and Time-Homogeneity Properties in Dempster-Shafer Random Walks. In: Ciucci, D., Couso, I., Medina, J., Slezak, D., Petturiti, D., Bouchon-Meunier, B. and Yager R.R. (Eds.), *Information Processing and Management of Uncertainty in Knowledge-Based Systems,* vol. 1601 of *Communications in Computer and Information Science*, Springer: 784–797.

Cinfrignini, A., Petturiti, D. and Vantaggi, B. (2023). Envelopes of equivalent martingale measures and a generalized no-arbitrage principle in a finite setting. *Annals of Operations Research*, 321: 103–137.

Coletti, G., Petturiti, D. and Vantaggi, B. (2015). Rationality principles for preferences on belief functions. *Kybernetika*, 51(3): 486–507.

Coletti, G., Petturiti, D. and Vantaggi, B. (2016). Conditional belief functions as lower envelopes of conditional probabilities in a finite setting. *Information Sciences*, 339: 64–84.

Coletti, G., Petturiti, D. and Vantaggi, B. (2020). A Dutch book coherence condition for conditional completely alternating Choquet expectations. *Bollettino dell'Unione Matematica Italiana*, 13(4): 585–593.

Coletti, G. and Vantaggi, B. (2008). A view on conditional measures through local representability of binary relations. *International Journal of Approximate Reasoning*, 47(3): 268–283.

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

Cuzzolin, F. (2021). *The Geometry of Uncertainty.* Artificial Intelligence: Foundations, Theory, and Algorithms. Springer.

de Cooman, G., De Bock, J. and Lopatatzidis, S. (2016). Imprecise stochastic processes in discrete time: global models, imprecise Markov chains, and ergodic theorems. *International Journal of Approximate Reasoning*, 76: 18–46.

de Cooman, G., Hermans, F. and Quaeghebeur, E. (2009). Imprecise Markov chains and their limit behaviour. *Probability in the Engineering and Informational Sciences*, 23(4): 597–635.

de Cooman, G., Troffaes, M. and Miranda, E. (2008). n-Monotone exact functionals. *Journal of Mathematical Analysis and Applications*, 347(1): 143-156.

Delbaen, F. and Schachermayer, W. (2006). *The mathematics of arbitrage.* Springer, Berlin.

Dempster, A. (1967). Upper and Lower Probabilities Induced by a Multivalued Mapping. *Annals of Mathematical Statistics*, 38(2): 325–339.

Denneberg, D. (1994). Conditioning (updating) non-additive measures. *Annals of Operations Research*, 52: 21–42.

Denœux, T. (2006). Constructing belief functions from sample data using multinomial confidence regions. *International Journal of Approximate Reasoning*, 42(3): 228–252.

Destercke, S. and Dubois, D. (2014). Special cases. In: Augustin, T., Coolen, F.P.A., de Cooman, G. and Troffaes, M.C.M. (Eds.), *Introduction to Imprecise Probabilities.* Wiley.

Dubins, L.E. (1975). Finitely additive conditional probabilities, conglomerability and disintegrations. *Annals of Probability*, 3(1): 89–99.

Dubois, D. and Prade, H. (1990). Updating with Belief Functions, Ordinal Conditioning Functions and Possibility Measures. *Proceedings of the Sixth Conference on Uncertainty and Artificial*, Cambridge: 307–315.

Dybvig, P.H. and Ross, S. (1989). Arbitrage. In: Eatwell, J., Milgate, M. and Newman, P. (Eds.), *Finance.* Palgrave Macmillan.

Edirsinghe, C., Naik, V. and Uppal, R. (1993). Optimal Replication of Options with Transactions Costs and Trading Restrictions. *The Journal of Financial and Quantitative Analysis*, 28(1): 117–138.

El Karoui, N. and Quenez, M.C. (1995). Dynamic Programming and Pricing of Contingent Claims in an Incomplete Market. *SIAM Journal on Control and Optimization*, 33(1): 29–66.

Ellsberg, D. (1961). Risk, Ambiguity, and the Savage Axioms. *The Quarterly Journal of Economics*, 75(4): 643–669.

Epstein, L.G. and Schneider, M. (2007). Learning Under Ambiguity. *The Review of Economic Studies*, 74(4): 1275–1303.

Epstein, L.G. and Schneider, M. (2008). Ambiguity, Information Quality, and Asset Pricing. *The Journal of Finance*, 63(1): 197–228.

Epstein, L.G. and Schneider, M. (2010). Ambiguity and Asset Markets. *Annual Review of Financial Economics, Annual Reviews*, 2(1): 315–346.

Fagin, R. and Halpern, J.Y. (1990). A new approach to updating beliefs. *Proceedings of the Sixth Conference on Uncertainty in Artificial Intelligence*, Cambridge: 347–374.

Fagin, R. and Halpern, J.Y. (1991). Uncertainty, belief, and probability. *Computational Intelligence*, 7: 160–173.

Faigle, U., Kern, W. and Still, G. (2002). *Algorithmic principle of mathematical programming.* (Texts in the Mathematical Sciences, 24), Springer.

Feng, D. and Nguyen H.T. (2007). Choquet weak convergence of capacity functionals of random sets. *Information Sciences*, 177(16): 3239–3250.

Garman, M.B. and Ohlson, J.A. (1980). Valuation of risky assets in arbitrage-free economies with transactions costs. *Journal of Financial Economics*, 9: 271–280.

Gilboa, I. and Marinacci, M. (2016). Ambiguity and the Bayesian Paradigm. In: Arló-Costa, H., Hendricks, V. and van Benthem. J. (Eds.). *Readings in Formal Epistemology.* Springer Graduate Texts in Philosophy, 1. Springer, Cham.

Gilboa, I. and Schmeidler, D. (1989). Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics*, 18(2): 141–153.

Gilboa, I. and Schmeidler, D. (1993). Updating ambiguous beliefs. *Journal of Economic Theory*, 59: 39–49.

Gilboa, I. and Schmeidler, D. (1994). Additive representations of non-additive measures and the Choquet integral. *Annals of Operations Research*, 52(1): 43–65.

Gilio, A. and Sanfilippo, G. (2011). Coherent conditional probabilities and proper scoring rules. *ISIPTA 2011 - Proceedings of the 7th International Symposium on Imprecise Probability: Theories and Applications.*

Gould, J.P. and Galai, D. (1974). Transaction Costs and the Relationship between Put and Call Prices. *Journal of Financial Economics*, 1(2): 105–129.

Grabish, M. (2016). *Set Functions, Games and Capacities in Decision Marking.* Springer.

Grimmett, G. and Stirzaker, D. (2020). *Probability and Random Processes.* Oxford University Press.

Harrison, J.M. and Kreps, D.M. (1979). Martingales and arbitrage in multiperiod securities markets. *Journal of Economic Theory*, 20(3): 381–408.

Harrison, J.M. and Pliska, S.R. (1980). Martingales and stochastic integrals in the theory of continuout trading. *Stochastic Processes and their Applications*, 11(3): 215–260.

Hartfiel, D.J. (1998). *Markov Set-Chains.* Springer-Verlag, Berlin.

Hermans, F. and Škulj, D. (2014). Stochastic processes. In: Augustin, T., Coolen, F.P.A., de Cooman, G. and Troffaes, M.C.M. (Eds.), *Introduction to Imprecise Probabilities.* Wiley.

Huber, P.J. (1981). *Robust Statistics.* Wiley, New York.

Hull, J.C. (2014). *Option, Futures and Other derivatives.* (9th edition). Pearson.

Jaffray, J.Y. (1989). Coherent bets under partially resolving uncertainty and belief functions. *Theory and Decision*, 26(2): 99–105.

Jaffray, J.Y. (2008). Bayesian Updating and Belief Functions. In: Yager, R.R. and Liu, L., *Classic Works of the Dempster-Shafer Theory of Belief Functions.* Studies in Fuzziness and Soft Computing, 219. Springer, Berlin, Heidelberg.

Jiang, W. (2018). A correlation coefficient for belief functions. *International Journal of Approximate Reasoning*, 103: 94–106.

Jouini, E. (2000). Price functionals with bid–ask spreads: an axiomatic approach. *Journal of Mathematical Economics*, 34(4): 547–558.

Jouini, E. and Kallal, L. (1995a). Martingales and Arbitrage in Securities Markets with Transaction Costs. *Journal of Economic Theory*, 66: 178–197.

Jouini, E. and Kallal, L. (1995b). Arbitrage in securities markets with short-sales constraints. *Mathematical Finance*, 5(3): 197–232.

Kast, R., Lapied, A. and Roubaud, D. (2014). Modelling under ambiguity with dinamically consistent Choquet random walks and Choquet-Brownian motions. *Economic Modelling*, 38: 495–503.

Klemkosky, R.C. and Resnick, B.G. (1979). Put-Call Parity and Market Efficiency. *Journal of Finance*, 34: 1141–1155.

Klibanoff, P., Marinacci, M. and Mukerji, S. (2009). Recursive smooth ambiguity preferences. *Journal of Economic Theory*, 144(3): 930–976.

Kozine, I.O. and Utkin, L.V. (2002). Interval-Valued finite markov chains. *Reliable Computing*, 8: 97–113.

Krak, T., T'Joens, N. and De Bock, J. (2019). Hitting Times and Probabilities for Imprecise Markov Chains. *Proceedings of Machine Learning Reserach*, 103: 265–275.

Liu, L. and Yager, R.R. (2008). *Classic works of the Dempster-Shafer Theory of Belief Functions: An Introduction.* Studies in Fuzziness and Soft Computing, Volume 219. Springer.

Melnikov, A.V. (1999). *Financial Markets: Stochastic Analysis and the Pricing of Derivative Securities.* Translations of Mathematical Monographs, American Mathematical Society.

Merton, R.C. (1990). *Continuous-time finance.* Blackwell.

Miranda, E., Montes, I. and Presa, A. (2022) Inner Approximations of Credal Sets by Non-additive Measures. In: Ciucci, D., Couso, I., Medina, J., Slezak, D., Petturiti, D., Bouchon-Meunier, B. and Yager R.R. (Eds.), *Information Processing and Management of Uncertainty in Knowledge-Based Systems,* vol. 1601 of *Communications in Computer and Information Science*, Springer: 743–756.

Miranda, E., Montes, I. and Vicig, P. (2021). On the selection of an optimal outer approximation of a coherent lower probability. *Fuzzy Sets Systems*, 424: 1–36.

Miyahara, Y. (1995). Canonical Martingale Measures of Incomplete Assets Markets. *Probability Theory and Mathematical Statistics*: Proceedings of the Seventh Japan-Russia Symposium, Tokyo.

Montes, I. and Destercke, S. (2017). On extreme points of p-boxes and belief functions. *Annals of Mathematics and Artificial Intelligence*, 81: 405–428.

Montes, I., Miranda, E. and Vicig, P. (2018). 2-Monotone outer approximations of coherent lower probabilities. *International Journal of Approximate Reasoning*, 101: 181–205.

Montes, I., Miranda, E. and Vicig, P. (2019). Outer approximating coherent lower probabilities with belief functions. *International Journal of Approximate Reasoning*, 110: 1–30.

Moral, S. (2018). Discounting Imprecise Probabilities. In: Gil, E., Gil, E., Gil, J. and Gil, M. (Eds.), *The Mathematics of the Uncertain*, vol. 142 of *Studies in Systems, Decision and Control*, Springer: 685–697.

Munk, C. (2013). *Financial Asset Pricing Theory.* Oxford University Press.

Nendel, M. (2021). Markov chains under nonlinear expectation. *Mathematical Finance*, 31: 474–507.

Palmer, K. (2001). A note on the Boyle-Vorst discrete-time option pricing model with transactions costs. *Mathematical Finance*, 11(3): 357–363.

Pascucci, A. and Runggaldier, W.J. (2011). *Financial Mathematics. Theory and problems for multi-period models.* Springer.

Pennesi, D. (2018). Asset prices in an ambiguous economy. *Mathematics and Financial Economics*, 12: 55–73.

Perrakis, S. and Lefoll, J. (1997). Derivative Asset Pricing with Transaction Costs: An Extension. *Computational Economics*, 10: 359–376.

Petturiti, D. and Vantaggi, B. (2020). Modeling agent's conditional preferences under objective ambiguity in Dempster-Shafer theory. *International Journal of Approximate Reasoning*, 119: 151–176.

Petturiti, D. and Vantaggi, B. (2022). Conditional decisions under objective and subjective ambiguity in Dempster-Shafer theory. *Fuzzy Sets and Systems*, 447: 155–181.

Petturiti, D. and Vantaggi, B. (2023). How to assess coherent beliefs: A comparison of different notions of coherence in Dempster-Shafer theory of evidence. *Reflections on the Foundations of Probability and Statistics: Essays in Honor of Teddy Seidenfeld*, Theory and Decision Library A, 54, DOI: 10.1007/978-3-031-15436-2_8.

Planchet, B. (1989). Credibility and conditioning. *Journal of Theoretical Probability*, 2(3): 289–299.

Pliska, S.R. (1997). *Introduction to mathematical finance: Discrete time models.* Blackwell, Malden, MA.

Predd, J.B., Seiringer, R., Lieb, E.H., Osherson, D.N., Poor, H.V. and Kulkarni, S.R. (2009). Probabilistic Coherence and Proper Scoring Rules. *IEEE Transactions on Information Theory*, 55(10): 4786–4792.

Reinganum, M.R. (1981a). Misspecification of capital asset pricing: Empirical anomalies based on earnings yields and market values. *Journal of Financial Economics*, 9: 19–46.

Reinganum, M.R. (1981b). The arbitrage pricing theory: Some empirical evidence. *Journal of Finance*, 36(2): 313–320.

Rényi, A. (1956). On conditional probability spaces generated by a dimensionally ordered set of measures. *Theory of Probability and its Applications*, 1(1): 55–64.

Ross, S.M. (2019). *Introduction to Probability Models.* (13th ed.) Academic Press, San Diego, CA, USA.

Roux, A. (2011). The fundamental theorem of asset pricing in the presence of bid-ask and interest rate spreads. *Journal of Mathematical Economics*, 47: 159–163.

Runggaldier, W.J. (2006). Portfolio optimization in discrete time. *Accademia delle scienze dell'Istituto di Bologna.*

Savage, L.J. (1954). *The Foundations of Statistics.* New York, Wiley.

Schervish, M.J., Seidenfeld, T. and Kadane, J.B. (2008). The fundamental theorems of prevision and asset pricing. *International Journal of Approximate Reasoning*, 49: 148–158.

Schmeidler, D. (1986). Integral representation without additivity. *Proceedings of the American Mathematical Society*, 97(2): 255–261.

Shafer, G. (1976a). *A Mathematical Theory of Evidence*. Princeton University Press, Princeton, NJ.

Shafer, G. (1976b). A theory of statistical evidence. In: Harper, W.L. and Hooker, C.A., *Foundation of probability theory, statistical inference, and statistical theories of science.* The University of Western Ontario Series in Philosophy of Science, 6b. Springer, Dordrecht.

Shafer, G. (1990). Perspectives on the Theory and Practice of Belief Functions. *International Journal of Approximate Reasoning*, 4 (5-6): 323–362.

Shafer, G. and Vovk, V. (2001). *Probability and finance: it's only a game*. Wiley.

Shafer, G. and Vovk, V. (2017). Towards a probability-free theory of continuous martingales. *arXiv:1703.08715*

Shafer, G. and Vovk, V. (2019). *Game-Theoretic Foundations for Probability and Finance*. Wiley.

Shreve, S. (1996). *Stochastic Calculus and Finance*. Springer.

Škulj, D. (2006). Finite Discrete Time Markov Chains with Interval Probabilities. In: Lawry, J., Miranda, E., Bugarin, A., Li, S., Gil, M.A., Grzegorzewski, P. and Hyrniewicz, O. (Eds.), *Soft Methods for Integrated Uncertainty Modelling.* Advances in Soft Computing, 37. Springer, Berlin, Heidelberg.

Škulj, D. (2009). Discrete time Markov chains with interval probabilities. *International Journal of Approximate Reasoning*, 50: 1314–1329.

Škulj, D. (2016). Random walks on graphs with interval weights and precise marginals. *International Journal of Approximate Reasoning*, 73: 76–86.

Sternberg, J. (1984). A Re-examination of Put-Call Parity of Index futures. *The Journal of Futures Markets*, 14(1): 79–101.

Suppes, P. and Zanotti, M. (1977). On Using Random Relations to Generate Upper and Lower Probabilities. *Synthese*, 36(4): 427–440.

T'Joens, N., De Bock, J. and de Cooman, G. (2021). A particular upper expectation as a global belief model for discrete-time finite-state uncertain processes. *International Journal of Approximate Reasoning*, 131: 30–55.

Tanner, E.J. and Levis, A.K. (1971). The Determinants of the Difference Between Bid and Ask Prices on Government Bonds. *The Journal of Business*, 44(4): 375–379.

Taylor, J.C. (1997). *An Introduction to Measure and Probability.* Springer.

Troffaes, M.C.M. and de Cooman, G. (2014). *Lower previsions.* Wiley.

Vasilev, I and Melnikov, A. (2021). On Market Completions Approach to Option Pricing. *Review of Business and Economics Studies*, 9: 77–93.

Walley, P. (1982). *Coherent lower (and upper) probabilities.* Tech. Rep. 22, Department of Statistics, University of Warwick, UK.

Walley, P. (1991). *Statistical Reasoning with Imprecise Probabilities.* Chapman and Hall, London.

Weichselberger, K. (2000). The theory of interval-probability as a unifying concept for uncertainty. *International Journal of Approximate Reasoning*, 24(2–3): 149–170.

Williams, P.M. (2007). Notes on conditional previsions. *International Journal of Approximate Reasoning*, 44(3): 366–383.

Yager, R.R. (1983). Entropy and specificity in a mathematical theory of evidence. *International Journal of General Systems*, 9(4): 249–260.

Yu, C. and Arasta, F. (1994). On Conditional Belief Functions. *International Journal of Approximate Reasoning*, 10: 155–172.