Swiss Finance Institute
Research Paper Series N°15-05

Economics-Based Financial Bubbles (and why they imply strict local martingales)

Martin HERDEGEN
ETH Zürich

Martin SCHWEIZER
ETH Zürich and Swiss Finance Institute
Economics-based financial bubbles
(and why they imply strict local martingales)

Martin Herdegen, Martin Schweizer*
ETH Zürich, Mathematik, HG G51.2,
Rämistrasse 101, CH–8092 Zürich, Switzerland
and*
Swiss Finance Institute, Walchestrasse 9, CH–8006 Zürich, Switzerland
{martin.herdegen,martin.schweizer}@math.ethz.ch

This version: February 18, 2015

Abstract
We introduce a new and numéraire-independent approach for defining and analysing financial bubbles in general, incomplete markets. We define our concepts in an economically motivated way using only primal quantities like assets and trading strategies. We then derive dual characterisations involving numéraires and martingale measures and show that a market is (interesting) bubbly in our sense if and only if all possible valuation measures for all possible discounted asset prices always lead to strict local martingales. In contrast to other approaches for bubble definitions in incomplete markets, our notion of a bubble is robust in the sense that it does not depend on the choice of a particular risk-neutral measure. We illustrate our results and concepts by many explicit and concrete examples; these include an incomplete market which is (interesting) bubbly, an incomplete market where one valuation measure sees a bubble while a second does not, and a natural setup where bubble birth occurs endogenously.

JEL classification: G10, C60
Keywords: bubble, incomplete financial market, fundamental value, strict local martingale, numéraire, viability, efficiency, no dominance

1 Introduction
This paper develops a new, economics-based approach for defining and analysing financial bubbles in a numéraire-independent paradigm. Let us first discuss some basic ideas.

The vast and diverse literature on bubbles is impossible to survey here, even only approximately. The Encyclopedia of Quantitative Finance has a 15-page entry “Bubbles and

*corresponding author
crashes” [32], with a list of over 100 references. A recent survey by Scherbina/Schlusche [50] puts more emphasis on behavioural models and rational models with frictions, and also provides a brief overview on the history of bubbles. The books of Brunnermeier [2] or Shiller [51] are often quoted as early classics; and the recent paper “A mathematical theory of financial bubbles” by Protter [46] also contains around 160 references plus some discussions of literature. At this point, we just recall some basic concepts.

In financial economics, the standard description of a bubble says that this is (linked to) an asset whose market value differs from (and usually exceeds) its fundamental value. A bit more formally, let us describe a dividend-paying asset \((\Delta, Y)\) by its cumulative dividend process \(\Delta = (\Delta_t)\) and its ex-dividend price process \(Y = (Y_t)\), both in the same fixed currency units. If \(Y_t^*\) denotes the asset’s (undiscounted) fundamental value at time \(t\), then \(Y_t^* \neq Y_t\) (or \(Y_t^* < Y_t\)) means that the asset has a bubble, and the difference \(Y_t - Y_t^*\) is usually called the (size of the) bubble or the bubble component of the asset.

The main difficulty with this very natural approach is that the notion of a “fundamental value” needs to be defined. There are different ideas for this; we discuss them in detail in Section 7, but have to mention them briefly here as well. One main school uses as (discounted) fundamental value the risk-neutral value of discounted future payments; this raises the question of the choice of risk-neutral measure. The other main school uses as fundamental value the superreplication cost of the asset; this brings up, in a more subtle and hidden way, a dependence on the chosen numéraire via the class of trading strategies allowed for doing superreplication. We follow the second method, but make sure that we a priori keep track of, and as far as possible eliminate, the dependence on the numéraire.

One key feature of our approach is that all our definitions are economically motivated and use only primal quantities like assets and trading strategies. Dual objects like numéraires and martingale measures also appear; but we employ them only in a second step to give dual characterisations of the notions introduced on the primal side. In particular, we show (Theorem 5.5) that strict local martingale measures arise naturally in the context of modelling financial bubbles. Moreover, our concept of bubbles also does not depend (neither a priori nor a posteriori) on a particular choice of a risk-neutral measure; in a sense made precise below, we therefore have a robust notion of bubbles. Last but not least, we illustrate our results by many concrete and explicit examples; these include an incomplete market which is (interesting) bubbly in our sense (Example 6.3), an incomplete market where one valuation measure sees a bubble while a second does not (Example 6.4), and a natural setup where bubble birth occurs endogenously (Example 3.7).

The paper is structured as follows. After explaining the main concepts of static and dynamic viability and efficiency in Section 2, we illustrate them by several examples in Section 3 before deriving their dual characterisations in Sections 4 and 5. We provide explicit examples of bubbly markets in Section 6, and finally compare our definitions and results to the existing literature on bubbles in Section 7.

1.1 Probabilistic setup and notation

We work throughout on an filtered probability space \((\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq T}, \mathbb{P})\) satisfying the usual conditions of right-continuity and completeness, where \(T > 0\) denotes a fixed and finite time horizon. We assume that \(\mathcal{F}_0\) is \(\mathbb{P}\)-trivial.
We call $\mathcal{T}_{[0,T]}$ the set of all stopping times with values in $[0,T]$. For $\sigma \in \mathcal{T}_{[0,T]}$, set $\mathcal{T}_{[\sigma,T]} := \{ \tau \in \mathcal{T}_{[0,T]} : \tau \geq \sigma \}$. For $\tau \in \mathcal{T}_{[0,T]}$, we denote by $L^0_+ (\mathcal{F}_\tau), \tilde{L}^0_+ (\mathcal{F}_\tau), L^0_{++} (\mathcal{F}_\tau)$ the set of all $\mathcal{F}_\tau$-measurable random variables taking $\mathbb{P}$-a.s. values in $[0,\infty), [0,\infty), (0,\infty)$, respectively. Finally, we denote by $e^i = (0,\ldots,0,1,0,\ldots,0)$ for $i = 1,\ldots,N$ the $i$-th unit vector in $\mathbb{R}^N$ and set $1 = \sum_{i=1}^N e^i = (1,\ldots,1) \in \mathbb{R}^N$.

A product-measurable process $\xi = (\xi_t)_{t \in [0,T]}$ is predictable on $[\sigma,T]$ if the random variable $\xi_\sigma$ is $\mathcal{F}_\sigma$-measurable and the process $\xi 1_{[\sigma,T]}$ is predictable. So if $\xi$ is predictable on $[\sigma,T]$ and $A \in \mathcal{F}_\sigma$, also $\xi 1_A$ is predictable on $[\sigma,T]$. For an $\mathbb{R}^N$-valued semimartingale $X = (X^1_t,\ldots,X^N_t)_{t \in [0,T]}$ and $\sigma \in \mathcal{T}_{[0,T]}$, we denote by $L_\sigma (X)$ the set of all $\mathbb{R}^N$-valued processes $\zeta = (\zeta^1_t,\ldots,\zeta^N_t)_{t \in [0,T]}$ which are predictable on $[\sigma,T]$ and for which the stochastic integral process $\int_\sigma^1 \zeta_s \cdot dX_s := \int_{[0,t]} \zeta_s 1_{[\sigma,t]}(s) \, dX_s, \, 0 \leq t \leq T$, is defined in the sense of $N$-dimensional stochastic integration (see [28, Section III.6] for details).

## 2 Main concepts

Throughout this paper, we consider a financial market with $N > 1$ assets and denote by $\tilde{S} = (\tilde{S}^1_t,\ldots,\tilde{S}^N_t)_{t \in [0,T]}$ the assets’ price process in some fixed but not specified currency unit. This unit may be tradable (e.g. in the form of a bank account) or not; we explicitly avoid assuming that one of the assets is constant 1, or that there exists a “bank account” $\tilde{S}^0$ in the background. All we initially impose is that the process $\tilde{S}$ is $\mathbb{R}^N$-valued, adapted and RCLL, that $\tilde{S}^i \geq 0$ $\mathbb{P}$-a.s. for each $i$, since we have primary assets in mind, and that the financial market is nondegenerate with $\tilde{S} \cdot 1$ strictly positive, meaning that

$$\inf_{t \in [0,T]} \tilde{S}_t \cdot 1 = \inf_{t \in [0,T]} \sum_{i=1}^N \tilde{S}^i_t > 0 \quad \mathbb{P}\text{-a.s.} \quad (2.1)$$

This excludes the case where all assets default and we are left with a nonexistent market.

It is folklore in mathematical finance that in a reasonable financial market, relative prices should be semimartingales after some suitable discounting; see e.g. Kardaras/Platen [34] and the references therein. To formalise this, we introduce the set $\tilde{\mathcal{D}}$ of one-dimensional adapted RCLL processes $\tilde{D} = (\tilde{D}_t)_{t \in [0,T]}$ with

$$\inf_{t \in [0,T]} \tilde{D}_t > 0 \quad \mathbb{P}\text{-a.s.} \quad (2.2)$$

and call the elements of $\tilde{\mathcal{D}}$ generalised exchange rate processes. We assume that

there exists some $\tilde{D} \in \tilde{\mathcal{D}}$ such that $\tilde{D}\tilde{S}$ is a semimartingale \quad (2.3)

and choose and fix one such $\tilde{D}$ and the corresponding process $S := \tilde{D}\tilde{S}$. We also call $S$ a semimartingale representative of the market described by $\tilde{S}$.

It is economically clear that all prices are relative and that the basic qualitative properties of a model should not depend on the chosen currency unit. To make this precise, we call a process $\tilde{S}'$ economically equivalent to $\tilde{S}$ if $\tilde{S}'$ is also $\mathbb{R}^N$-valued, adapted and RCLL, and if $\tilde{S}' = \tilde{D}'\tilde{S}$ for some $\tilde{D}' \in \tilde{\mathcal{D}}$. In other words, two processes are economically equivalent if they describe the same assets in possibly different currency units.

Our first simple result shows that our modelling approach does not depend on the initial choice of $S$ and that it has nice semimartingale properties, in the following sense.
Lemma 2.1. Suppose $\tilde{S}$ and $\tilde{S}'$ are economically equivalent. If $\tilde{S}$ satisfies (2.1) or (2.3), then so does $\tilde{S}'$. If $\tilde{S}$ satisfies both (2.1) and (2.3) and we choose a semimartingale representative $S = \tilde{D}\tilde{S}$, then each semimartingale representative $S' = \tilde{D}'\tilde{S}'$ is economically equivalent to $S$ with an exchange rate process $D \in \mathcal{D}$ which is even a semimartingale.

Proof. Since $\tilde{D}' > 0$ in the sense of (2.2), (2.1) directly transfers from $\tilde{S}$ to $\tilde{S}' = \tilde{D}'\tilde{S}$. From (2.3) for $\tilde{S}$, we obtain $S = \tilde{D}\tilde{S}$ for some semimartingale $S$ and some $\tilde{D} \in \tilde{\mathcal{D}}$. So $\tilde{S} = \tilde{D}(\tilde{S}'/\tilde{D}') = (\tilde{D}/\tilde{D}')\tilde{S}'$ is a semimartingale and $\tilde{D}/\tilde{D}'$ is in $\tilde{\mathcal{D}}$, and we see that $\tilde{S}'$ also satisfies (2.3). If $S' = \tilde{D}'\tilde{S}'$ is a semimartingale, we can use $\tilde{S}' = \tilde{D}'\tilde{S}$ to write $S' = DS$ with $D := D'(\tilde{D}'/\tilde{D})$ which is clearly in $\tilde{\mathcal{D}}$. But $S = \tilde{D}\tilde{S}$ and $S' = \tilde{D}'\tilde{S}'$ both also satisfy (2.1), and so we can write $D = (S' \cdot 1)/(S \cdot 1)$ to see that $D$ is also a semimartingale. □

In the sequel, we always assume that (2.1) and (2.3) are satisfied, and we choose a semimartingale representative $S$. All other semimartingale representatives are then economically equivalent to $S$ with a semimartingale exchange rate process, and we introduce the set of exchange rate processes,

$$\mathcal{D} := \tilde{\mathcal{D}} \cap \{\text{semimartingales}\} = \{\text{all one-dimensional semimartingales } D = (D_t)_{t \in [0,T]} \text{ with } \inf_{t \in [0,T]} D_t > 0\},$$

and the market generated by $S$, which is

$$S := \{S' = DS : D \in \mathcal{D}\}.$$

The key difference between $S$ and $\tilde{S}$ is that $S$ is a semimartingale, and we exploit this when we formalise trading and self-financing strategies with the help of stochastic integrals. Up to a change of currency unit, however, $S$ and $\tilde{S}$ agree; so we can view the choice of working with $S$ as merely dictated by convenience, and we can always rewrite everything back into the units of $\tilde{S}$ if that is preferred for some reason; see Remark 2.5 below for more details.

Example 2.2 (Classic setup of mathematical finance). One particular case is what we call the classic setup of mathematical finance. Suppose there is one asset which has for $\mathbb{P}$-almost all $\omega$ a positive price. (More precisely, we need $\mathbb{P}[\inf_{t \in [0,T]} \tilde{S}_t^k > 0] = 1$ for some $k$, so that $B := \tilde{S}^k \in \tilde{\mathcal{D}}$ is a generalised exchange rate process.) Then we can express all other assets in units of that special asset by defining $X^i := \tilde{S}^i/\tilde{S}^k$ and then relabel the assets; we call that particular asset $k$ now asset 0 or bank account, and we call the other $d := N - 1$ assets the risky assets, discounted by the bank account. For later use, we also introduce the vector process $Y := (\tilde{S}^1, \ldots, \tilde{S}^{k-1}, \tilde{S}^{k+1}, \ldots, \tilde{S}^N)$ so that $\tilde{S}$ can be identified both with $(B, Y)$ and with $(1, X)$ given $B$. Then we have $N = d + 1$ basic assets; but they are not symmetric because one of them is a riskless bank account which can never reach the value 0. Moreover, if there are several assets like $\tilde{S}^k$, the choice of the one we use for discounting is arbitrary. So if we define concepts in terms of $X = Y/B$, they depend implicitly on the chosen discounting process $\tilde{S}^k$, and it can become difficult to keep track of this dependence all the time.

The vast majority of papers in mathematical finance—with the obvious exception in the literature on interest rate modelling—works with the end result of the above setup. Usually, papers start with an $\mathbb{R}^d$-valued process $X$ and call this the (discounted) price
process of \( d \) risky assets. Almost without exception, it is also assumed (but very often not mentioned explicitly) that there is in addition to \( X \) a riskless bank account whose price is identically 1—and this assumption is exploited in the standard problem formulations. (Most papers also assume that \( X \) is a semimartingale, which corresponds to our choice of a semimartingale representative \( S \).)

As one can see, the classic setup is intrinsically asymmetric. This hides or obscures a number of important phenomena, and so we want to start with a symmetric treatment of all assets. Since we make no assumptions on \( \tilde{D} \) in (2.2) except strict positivity, all our results include the classic setup with nonnegative prices; but they do not exploit its assumptions and asymmetry, and hence they are both more general and in our view more natural. The simplest example of a model which cannot be formulated in the classic setup is one with two assets \( (N = 2) \); they are both nonnegative, but both can default, i.e. become 0. One of them hits 0 at some (maybe random) time on a set \( A \) only; the other hits 0 on \( A^c \) only. If \( 0 < P[A] < 1 \), this cannot be put into the form of the classic setup. For a more detailed and intuitive formulation, see Example 3.1 in Section 3 below.

**Remark 2.3.** One can of course argue in the above example that introducing a third asset of the form \( S^3 = \alpha S^1 + (1 - \alpha)S^2 \) with \( \alpha \in (0,1) \) would lead us back into the classic setup without changing the market, since we have the same trading opportunities. However, this easy way out is an ad hoc fix, and also raises the question how the resulting classic setup depends perhaps on the choice of \( \alpha \). Rather than trying to find a case-by-case approach, we prefer to deal with (2.1) and (2.3) in a general and systematic way.

Now let us return to our basic model. We want to describe (frictionless) continuous trading and work with self-financing strategies; so we need stochastic integrals, and therefore exploit below that \( S \) is a semimartingale. Again, this includes the classic setup.

One direct consequence of our symmetric formulation is as follows. If \( \tilde{D} \in \tilde{D} \) is any generalised exchange rate process, the process \( \tilde{S}^\prime := \tilde{D}' S \) describes the same market, but in a different currency unit. (Like the original currency unit, the new unit induced by the “exchange rate” \( \tilde{D}' \) may be available for trade or not.) So our initially chosen \( S \in S \) is just one, fixed, semimartingale representative of the market \( S \) generated by \( S \) or \( \tilde{S} \).

In the sequel, we only want to work with notions which are independent of the choice of a specific semimartingale representative \( S' \in S \) (or a particular currency unit). More precisely, a notion should hold for our fixed semimartingale representative \( S \) if and only if it holds for each \( S' \in S \); then we say that the notion holds for the market \( S \) and call it “numéraire-independent”. Wherever this “numéraire independence” is not directly clear from the context or the definitions, we shall make a comment or give an explanation.

### 2.1 Self-financing strategies, numéraires and strategy cones

In this section, we introduce trading strategies. This is almost standard, with small (but important) differences because we are not in the classic setup. Recall that \( S \in S \) is a fixed semimartingale representative of the market \( S \).

**Definition 2.4.** Fix a stopping time \( \sigma \in T_{[0,T]} \). A self-financing strategy (for \( S \)) on \([\sigma,T] \)
is an $N$-dimensional process $\vartheta$ which is predictable on $[\sigma, T]$, in $L_\sigma(S)$, and such that

$$V(\vartheta)(S) := \vartheta \cdot S = \vartheta_\sigma \cdot S_\sigma + \int_\sigma^T \vartheta_u \, dS_u \quad \text{on } [\sigma, T], \, \mathbb{P}\text{-a.s.} \quad (2.4)$$

We denote the space of all these strategies by $L^d_\sigma(S)$ or just $L^d_\sigma$, and we call $V(\vartheta)(S)$ the value process of $\vartheta$ (in the currency unit corresponding to $S$).

It is not immediately obvious but true that the above concept of a self-financing strategy is “numéraire-independent”. In fact, one can show for each $S' \in \mathcal{S}$ that if $\vartheta$ is in $L_\sigma(S)$ and satisfies (2.4), then $\vartheta$ is also in $L_\sigma(S')$ and satisfies (2.4) for $S'$ instead of $S$; see Herdegen [23, Lemma 2.8], and note that the proof for $\sigma \neq 0$ is verbatim the same as for $\sigma = 0$. In particular, writing $L^d_\sigma(S)$ and not $L^d_\sigma(S)$ is justified. Note that the value process of any strategy $\vartheta$ satisfies the “exchange rate consistency property”

$$V(\vartheta)(DS) = DV(\vartheta)(S) \quad \text{for every exchange rate process } D \in \mathcal{D}. \quad (2.5)$$

This means that when we change units from $S$ to $S' = DS$, the wealth from $\vartheta$ in new units is simply the old wealth multiplied by $D$, as it must be from basic financial intuition.

**Remark 2.5.** 1) Combining (2.5), (2.4) and the numéraire independence of (2.4) gives

$$V(\vartheta)(S') = V_\sigma(\vartheta)(S') + \frac{1}{D'} \int_\sigma^T \vartheta_u \, d(S'D')_u \quad \text{on } [\sigma, T], \, \mathbb{P}\text{-a.s.},$$

for all $\vartheta \in L^d_\sigma$ and semimartingale representatives $S' = S/D'$. This change-of-numéraire formula has appeared, among others, in El Karoui et al. [18] or Takaoka/Schweizer [54]. Since it follows from the definition (2.4) for any semimartingale representative $S'$, it is natural to extend it by definition to all other representatives $\tilde{S}$ as well. So if we want to work with self-financing strategies not for $S$, but the original (possibly non-semimartingale) $\tilde{S} = S/D$, we rewrite the self-financing condition (2.4) in the units corresponding to $\tilde{S}$ by multiplying everything by the exchange rate $D$ at the appropriate time, i.e., as

$$V(\vartheta)(\tilde{S}) := \vartheta \cdot \tilde{S} = \vartheta_\sigma \cdot \tilde{S}_\sigma + \frac{1}{D} \int_\sigma^T \vartheta_u \, d(\tilde{S}D)_u \quad \text{on } [\sigma, T], \, \mathbb{P}\text{-a.s.} \quad (2.6)$$

This avoids the need of defining stochastic integrals with respect to $\tilde{S}$ (which might even be impossible).

2) In the classic setup with $N = d + 1$, $S = (1, X)$ and discounted asset prices given by the $\mathbb{R}^d$-valued semimartingale $X$, self-financing strategies on $[\sigma, T]$ can be identified with pairs $(v_\sigma, \psi)$ of $\mathcal{F}_\sigma$-measurable random variables $v_\sigma$ and $\mathbb{R}^d$-valued predictable $X$-integrable processes $\psi$. Indeed, setting $v_\sigma := V_\sigma(\vartheta)(S)$ and using that asset 0 has a constant price of 1, we can write (2.4) for a strategy $\vartheta = (\eta, \psi)$ in $S = (1, X)$ as

$$\eta = V(\vartheta)(S) - \psi \cdot X = v_\sigma + \int_\sigma^T \vartheta_u \, dX_u - \psi \cdot X;$$

see e.g. Föllmer/Schied [17, Remark 5.8] or Elliott/Kopp [12, Lemma 2.2.1]. Since it is so familiar, this identification of $\vartheta$ with $(v_\sigma, \psi)$ is even done without mention in most papers. In our symmetric setup, such a simple identification is no longer possible; trading strategies must be treated as processes of dimension $N = d + 1$, and the self-financing condition (2.4) imposes a linear constraint on their coordinates.
Clearly, $L^\text{sf}_\sigma(S)$ is a vector space. It is also closed under multiplication with $\mathcal{F}_\sigma$-measurable random variables; so we can scale a strategy on $[\sigma, T]$ not only by a constant, but also by a random factor known at the beginning $\sigma$ of the time period on which we trade.

To avoid doubling phenomena, one usually considers sub-cones of $L^\text{sf}_\sigma$ for “allowed” trading. We first give the abstract definition.

**Definition 2.6.** For a stopping time $\sigma \in \mathcal{T}_{[0,T]}$, a strategy cone (for $S$) on $[\sigma,T]$ is a nonempty subset $\Gamma \subseteq L^\text{sf}_\sigma(S)$ with the properties

1) if $\vartheta^1, \vartheta^2 \in \Gamma$ and $c^1 \vartheta^1, c^2 \vartheta^2 \in L^0_+ (\mathcal{F}_\sigma)$, then $c^1 \vartheta^1 + c^2 \vartheta^2 \in \Gamma$;

2) if $(\vartheta^n)_{n \in \mathbb{N}}$ is a countable family in $\Gamma$ and $(A_n)_{n \in \mathbb{N}}$ a partition of $\Omega$ into $\mathcal{F}_\sigma$-measurable sets, then $\sum_{n=1}^{\infty} \mathbf{1}_{A_n} \vartheta^n \in \Gamma$.

A family of strategy cones $(\Gamma_\sigma)_{\sigma \in \mathcal{T}_{[0,T]}}$, where each $\Gamma_\sigma$ is a strategy cone on $[\sigma,T]$, is called time-consistent if $\Gamma_{\sigma_1} \subseteq \Gamma_{\sigma_2}$ for $\sigma_1 \leq \sigma_2$ in $\mathcal{T}_{[0,T]}$.

The simplest example of a strategy cone on $[\sigma,T]$ is $L^0_\sigma$ itself. Moreover, the family $(L^\text{sf}_\sigma)_{\sigma \in \mathcal{T}_{[0,T]}}$ is clearly time-consistent. The main example used in this paper is given by the class of undefaultable strategies introduced below in Definition 2.8.

If $\Gamma \subseteq L^\text{sf}_\sigma(S)$ is a strategy cone on $[\sigma,T]$, we set, for any norm $\| \cdot \|$ in $\mathbb{R}^N$,

$$b\Gamma := \left\{ \vartheta \in \Gamma : \sup_{(\omega,t) \in \Pi \times [0,T]} \| \vartheta \mathbf{1}_{[\vartheta,T]} \| \leq c_\sigma \text{ for some } c_\sigma \in L^0_+ (\mathcal{F}_\sigma) \right\},$$

$$h\Gamma := \left\{ \vartheta \in \Gamma : \vartheta \mathbf{1}_{[\sigma,T]} = \vartheta_\sigma \mathbf{1}_{[\sigma,T]} \right\}.$$ 

Clearly, $\{0\} \subseteq h\Gamma \subseteq b\Gamma \subseteq \Gamma$, and $h\Gamma$ and $b\Gamma$ are again strategy cones on $[\sigma,T]$. We call $\vartheta \in b\Gamma$ a bounded strategy in $\Gamma$ and $\vartheta \in h\Gamma$ an invest-and-keep strategy in $\Gamma$. Note that if the family $(\Gamma_\sigma)_{\sigma \in \mathcal{T}_{[0,T]}}$ is time-consistent, then $(b\Gamma_\sigma)_{\sigma \in \mathcal{T}_{[0,T]}}$ and $(h\Gamma_\sigma)_{\sigma \in \mathcal{T}_{[0,T]}}$ are so, too.

**Remark 2.7.** 1) An invest-and-keep strategy is the simplest and most naive strategy, requiring from the investor only one decision: He should decide when to start and how much of each asset he wants. Then he just waits until the end $T$ of the trading interval.

2) Calling strategies in $b\Gamma$ bounded may seem puzzling at first sight. But each $\vartheta \in b\Gamma$ is uniformly bounded on $[\sigma,T]$ by an $\mathcal{F}_\sigma$-measurable random variable $c_\sigma \in L^0_+ (\mathcal{F}_\sigma)$, and the latter play the role of “constants” on $[\sigma,T]$. (Recall that $L^\text{sf}_\sigma$ is closed under multiplication with elements of $L^0_+ (\mathcal{F}_\sigma)$, and we stipulate the cone structure in Definition 2.6 for elements of $L^0_+ (\mathcal{F}_\sigma)$.) In particular, for $\sigma = 0$, we recover the usual concept of a bounded strategy.

3) Our strategies are parametrised in numbers of units of assets (“shares”), not in wealth amounts or fractions of wealth. So for a bounded strategy, asset holdings are bounded but wealth need not be.

It is well known that to avoid undesirable phenomena in a financial market, one must exclude doubling-type strategies. The usual way to do that is to impose solvency constraints—strategies are allowable for trading only if their value processes are bounded below by some quantity. If this approach should not depend on a specific currency unit, the only possible choice for the lower bound is 0. This motivates the following definition.
Definition 2.8. Fix a stopping time \( \sigma \in T_{[0,T]} \). We call a strategy \( \vartheta \in L^f_\sigma(S) \) an undefaultable strategy on \([\sigma,T]\) and write \( \vartheta \in \mathcal{U}_\sigma(S) \) or just \( \vartheta \in \mathcal{U}_\sigma \) if
\[
V(\vartheta)(S) \geq 0 \quad \text{on } [\sigma,T], \ \mathbb{P}\text{-a.s.}
\]

The notion of undefaultable is clearly “numéraire-independent”. Moreover, each \( \mathcal{U}_\sigma \) is a strategy cone, and \( (\mathcal{U}_\sigma)_{\sigma \in T_{[0,T]}} \) is a time-consistent family.

The next concept we need is a “numéraire”. We first give the definition and then some comments. (In the classic setup, it is actually rare to find a precise definition of the term “numéraire”; the expression seems to be viewed as self-explaining.)

Definition 2.9. A numéraire strategy (for the market \( S \)) is a strategy \( \eta \in L^f(S) \) with
\[
\inf_{t \in [0,T]} V_t(\eta)(S) > 0 \quad \text{\( \mathbb{P}\)\,-a.s., i.e., } V(\eta)(S) \in \mathcal{D}.
\]
We call \( S \) a numéraire market if such an \( \eta \) exists.

Note that the above concept is “numéraire-independent” since \( V(\eta)(S) > 0 \) holds for some \( S \in \mathcal{S} \) if and only if it holds for all \( S' \in \mathcal{S} \), due to the exchange rate consistency (2.5). Note also that any numéraire strategy is automatically in \( \mathcal{U}_0(S) \).

By our standing nondegeneracy assumption (2.1), the market portfolio \( \eta^S := 1 \) of holding one unit of each asset is always a numéraire strategy; it even lies in \( h\mathcal{U}_0(S) \) and hence is bounded. Similarly, in a classic setup \( S = (B,Y) \) of \( N = d + 1 \) assets, where \( Y \) denotes \( d \) undiscounted “risky assets” and \( B \in \tilde{D} \) an undiscounted “bank account”, the invest-and-keep strategy \( e^1 \) of the “bank account” is a bounded numéraire strategy. For general \( \eta \), in the classic setup, one calls \( V(\eta)(S) \) (which equals \( B \) in the latter example) a “numéraire” or “tradable numéraire”. But doing that implies that we work in the currency unit corresponding to the particular representative \( S \), and such a dependence is just what we want to avoid. Thus we describe “numéraires” not by their wealth, but their asset holdings, which are numbers (of “shares”) and do not depend on any currency unit.

For each numéraire strategy \( \eta \), there exists a \( \mathbb{P}\)-a.s. unique numéraire representative \( S^{(\eta)} \in \mathcal{S} \) such that \( V(\eta)(S^{(\eta)}) \equiv 1 \). It is given explicitly by “\( V(\eta)\)-discounted prices”
\[
S^{(\eta)} := \frac{S}{V(\eta)(S)}.
\] (2.7)

Note that \( S^{(\eta)} \) is well defined because \( V(\eta) \) satisfies the exchange rate consistency property (2.5); this ensures that the right-hand side of (2.7) yields the same result for any other representative \( S' = DS \) of \( \mathcal{S} \). In the classic setup \( S = (B,Y) \) as above with a bank account \( B \) and \( \eta = e^1 \), (2.7) reduces to \( S^{(e^1)} = S/B = (1,X) \) as in Example 2.2.

2.2 Maximal strategies

Suppose we are given a class \( \Gamma \) of possible strategies. A strategy \( \vartheta \in \Gamma \) can be considered a “reasonable investment” from that class only if it cannot be directly improved by another strategy from the same class. More precisely, using strategies in \( \Gamma \) with the same (or a lower) initial investment should not allow one to create more wealth at time \( T \). It is natural to call such a strategy \( \vartheta \) maximal; see Remark 2.11 below for more comments.
Definition 2.10. Let $\sigma \in \mathcal{T}_{[0, T]}$ be a stopping time and $\Gamma$ a strategy cone on $[\sigma, T]$. A strategy $\vartheta \in \Gamma$ is weakly maximal for $\Gamma$ if there is no pair $(f, \bar{\vartheta}) \in (L_0^+(\mathcal{F}_T) \setminus \{0\}) \times \Gamma$ with

$$V_T(\bar{\vartheta})(S) \geq V_T(\vartheta)(S) + f \ \mathbb{P}\text{-a.s.} \quad \text{and} \quad V_\sigma(\bar{\vartheta})(S) \leq V_\sigma(\vartheta)(S) \ \mathbb{P}\text{-a.s.}$$

It is strongly maximal for $\Gamma$ if there is no $f \in L_0^+(\mathcal{F}_T) \setminus \{0\}$ such that for all $\varepsilon > 0$, there exists $\bar{\vartheta} \in \Gamma$ with

$$V_T(\bar{\vartheta})(S) \geq V_T(\vartheta)(S) + f \ \mathbb{P}\text{-a.s.} \quad \text{and} \quad V_\sigma(\bar{\vartheta})(S) \leq V_\sigma(\vartheta)(S) + \varepsilon \ \mathbb{P}\text{-a.s.} \quad (2.8)$$

Note above that $f$, which satisfies $f \geq 0 \ \mathbb{P}\text{-a.s.}$ and $\mathbb{P}[f > 0] > 0$, stands for the extra wealth at time $T$, on top of what we get from $\vartheta$, that we generate by $\bar{\vartheta}$. If $\vartheta$ is weakly maximal, no $\bar{\vartheta}$ achieves $f$ without increasing the initial capital at time $\sigma$. If $\vartheta$ is strongly maximal, the improvement is asymptotically impossible even if we are allowed a small but strictly positive increase of initial capital at $\sigma$. Both concepts are clearly “numéraire-independent”.

Remark 2.11. 1) The terminology “maximal strategy” goes back at least to Delbaen/Schachermayer [4, 6, 7]. However, their setting is different from ours so that maximality has a different meaning. More precisely, their papers are cast in the classic setup, use for $\Gamma$ the class $\mathcal{A}$ (which is not “numéraire-independent” in our sense) of so-called admissible strategies on $[0, T]$, and a priori impose an absence-of-arbitrage condition. In our terminology, a maximal strategy in the sense of e.g. [7] is then weakly maximal for $\mathcal{A}$.

2) Both above definitions of maximality are slightly different from those in Herdegen [23, Definitions 3.1 and 3.9]. In [23], a strategy is called maximal if it is maximal for $\Gamma$ on $[0, \sigma]$, for every stopping time $\sigma \in \mathcal{T}_{[0, T]}$. However, under a natural extra assumption on $\Gamma$ ([23, Definition 3.5]), a strategy which is (weakly or strongly) maximal in our sense for $\Gamma$ (i.e. on $[0, T]$) is also (weakly or strongly) maximal in the sense of [23]. The above assumption (which essentially amounts to predictable convexity) is in particular satisfied for the class $\mathcal{U}$ of undefaultable strategies. Finally, Corollary 4.11 below also shows that maximality in $\mathcal{U}_0$ (i.e. on $[0, T]$) is equivalent to maximality in each $\mathcal{U}_\sigma$ (i.e. on $[\sigma, T]$).

3) It is clear that strong implies weak maximality, and the converse does not hold in general; see Example 3.5 below in Section 3. But if the zero strategy 0 is strongly maximal for $\mathcal{U}_\sigma$, $\sigma \in \mathcal{T}_{[0, T]}$, then weak implies strong maximality for $\mathcal{U}_\sigma$; see Lemma 5.1 below.

2.3 Viability and efficiency criteria for markets

A financial market should behave in a reasonable way, and this should be reflected in the properties of its model description. Let us formalise this and then explain the intuition.

Definition 2.12. A market $\mathcal{S}$ is called

• statically viable if the zero strategy 0 is strongly maximal for $\mathcal{hU}_\sigma(\mathcal{S})$, for each $\sigma \in \mathcal{T}_{[0, T]}$.

• dynamically viable if the zero strategy 0 is strongly maximal for $\mathcal{U}_\sigma(\mathcal{S})$, for each $\sigma \in \mathcal{T}_{[0, T]}$. 

9
Static viability means that at every time $\sigma$, just doing nothing cannot be improved by a self-financing invest-and-keep strategy. Dynamic viability is even stronger—one cannot improve on inactivity by trading even if one trades continuously in time. Of course, in both cases, one must obey the constraint (from $\mathcal{U}_\sigma$) of keeping wealth nonnegative.

Dynamic viability by its definition implies static viability, but the converse is not true. Even in a finite-state discrete-time setup (with more than one time period), static is strictly weaker than dynamic viability; see Example 3.6 below. For finite discrete time, we show later in Proposition 4.2 that dynamic viability is equivalent to the classic no-arbitrage condition $\text{NA}$. In general, Corollary 4.11 below implies that a market $\mathcal{S}$ is dynamically viable if and only if the zero strategy is strongly maximal for $\mathcal{U}_0$; so it is enough to check maximality for the starting time 0 instead of all $\sigma \in T[0,T]$. Strong maximality of 0 in $\mathcal{U}_0$ has been coined numéraire-independent no-arbitrage (NINA) and analysed in detail in Herdegen [23]. From all these results, we see that dynamic viability can be viewed as a (weak and general) property of absence of arbitrage.

The next concept strengthens viability.

**Definition 2.13.** A market $\mathcal{S}$ is called

- **statically efficient** if each strategy $\vartheta \in \mathcal{hU}_\sigma(\mathcal{S})$ is strongly maximal for $\mathcal{hU}_\sigma(\mathcal{S})$, for each $\sigma \in T[0,T]$.

- **dynamically efficient** if each strategy $\vartheta \in \mathcal{hU}_\sigma(\mathcal{S})$ is strongly maximal for $\mathcal{U}_\sigma(\mathcal{S})$, for each $\sigma \in T[0,T]$.

Viability means that one cannot improve the zero strategy of doing nothing. An efficient market has even more structure—all invest-and-keep strategies are good in the sense that they cannot be improved, in a certain class, without risk or extra capital. It is clear from the definition that dynamic implies static efficiency, and like for viability, the converse is not true; this is also illustrated below in Example 3.6.

The connection between viability and efficiency is more subtle. Clearly efficiency (dynamic or static) implies viability (of the same kind). At first sight, one might expect the converse as well—why should it matter whether one improves zero or a general invest-and-keep strategy? However, there is a difference, and the reason is that one must look for improvements in the class of undefaultable strategies which is a cone, but not a linear space. If a strategy is to be improved and we take differences to construct something better, this leads us outside that cone in general—except of course if we subtract zero.

Interestingly and notably, the difference between efficiency and viability does not appear in finite discrete time. Proposition 4.2 below shows that dynamic efficiency is equivalent to dynamic viability in finite discrete time, and one can show (see [22, Lemma VIII.1.19] for a proof) that the two static concepts are also equivalent, for finite discrete time. This reflects the well-known fact that if one can achieve arbitrage in finite discrete time with a general strategy, one can also achieve arbitrage with an undefaultable strategy. However, this is specific to finite discrete time because the proof relies on backward induction; see for example Elliott/Kopp [12, Section 2.2].

In a market with infinitely many trading dates, things change. In continuous time, Example 6.1 shows that static/dynamic viability in general does not imply static/dynamic efficiency. So the next concepts are meaningful.
Definition 2.14. A market $S$ is called a *bubbly market* if it fails to be dynamically efficient. It is called an *interesting* bubbly market if it is in addition dynamically viable, and a *nontrivial* bubbly market if yet in addition, it is statically efficient.

The idea behind this definition is simple. A market $S$ is bubbly if some invest-and-keep strategy can be improved (approximately) by dynamic trading. Of course, this will happen if $S$ “allows arbitrage”, i.e. if $S$ is not dynamically viable; but this is not the really interesting situation of a bubbly market because $S$ is then already “too degenerate”. If we also want to consider valuation aspects, we typically use dynamic trading in the basic assets and semi-static trading in the additional instruments (e.g. options); so it makes sense to assume that such strategies cannot be improved, and therefore we also impose static efficiency for a nontrivial bubbly market. In the present paper, we only analyse interesting bubbly markets; nontrivial ones will be studied in forthcoming work.

Note that an interesting bubbly market can only appear in a model with infinitely many trading dates. In finite discrete time, the equivalence of dynamic viability and the no-arbitrage property NA (Proposition 4.2) implies that every bubbly market allows arbitrage and hence cannot be interestingly bubbly. We believe that this dichotomy is natural and some interesting phenomena inherently need an infinite set of trading dates.

Remark 2.15. 1) Throughout this paper, our setting has a last trading date; we either work in continuous time on the (right-closed) interval $[0,T]$ or in discrete time on $\{0,1,\ldots,T\}$ (then with $T \in \mathbb{N}$). We believe that results like those for $[0,T]$ can also be developed for trading dates in $[0,\infty)$ or in $\mathbb{N}_0 = \{0,1,2,\ldots\}$, but one must take some extra care as time goes to $\infty$. This is left for future research.

2) Note that in contrast to much of the existing work on bubbles, our definitions do not involve any strict local martingale property of some asset prices. We start instead with an economically compelling notion of a bubbly market, and then prove that there must be a close connection to strict local martingales. This is done below in Theorem 5.5.

3) The next section shows that our definition of a bubbly market is (almost) equivalent to the standard definition from financial economics; see Proposition 2.19 below. We discuss the connections to the literature in more detail in Section 7.

### Definition of bubbles via superreplication prices

As pointed out in the introduction, the standard approach in financial economics is to define bubbles by comparing market prices to fundamental values. We now show that if fundamental values are given by superreplication prices, at any time $t$, this gives the same concept of bubbles as introduced just above.

**Definition 2.16.** The *superreplication price* of asset $i \in \{1,\ldots,N\}$ at time $\tau \in T_{[0,T]}$ in the currency unit corresponding to $S$ is

$$
\pi_i^\tau(S) := \text{ess inf}\{v \in L^0_+(\mathcal{F}_\tau) : \exists \vartheta \in U_\tau \text{ with } V_T(\vartheta)(S) \geq S_T^\tau \text{ and } V_\tau(\vartheta)(S) \leq v \ P\text{-a.s.}\}
$$

**Definition 2.17.** The *fundamental value* of asset $i \in \{1,\ldots,N\}$ at time $\tau \in T_{[0,T]}$ in the currency unit corresponding to $S$ is defined by

$$
^*S^\tau_i := \pi_i^\tau(S).
$$
We say that $\mathcal{S}$ is an economically bubbly market if there exist $\tau \in \mathcal{T}_{[0,T]}$ and an asset $i \in \{1, \ldots, N\}$ such that for some $S \in \mathcal{S}$,
\[ P[\tilde{S}^i_\tau < S^i_\tau] > 0. \]

If we accept the idea that the fundamental value of an asset should be given by its superreplication price, the above definition clearly formalises the standard idea of a bubble from financial economics. However, it is more general than what one can find so far since we compare $S$ and $^*S$ not only at time 0. In particular, it may happen that $\mathcal{S}$ is economically bubbly, but has $S_0 = ^*S_0$, and in that sense, our definition of an economically bubbly market includes the possibility of “bubble birth”. We give an explicit example for such a market in Example 3.7 below and provide more discussion in Section 7.4 below.

Remark 2.18. 1) Since the invest-and-keep strategy $e^i$ of holding one unit of asset $i$ is in $\mathcal{U}$ and has $V_\tau(e^i)(S) = S^i_\tau$ and $V_\tau(e^\tau)(S) = S^i_\tau$, we clearly have $^*S^i_\tau = \pi^i_\tau(S) \leq S^i_\tau$. Moreover, the exchange rate consistency (2.5) implies that we also have
\[ ^*(DS)^i_\tau = \pi^i_\tau(DS) = D_\tau \pi^i_\tau(S) = D_\tau ^*S^i_\tau \quad \text{for all } D \in \mathcal{D}. \]
As a consequence, the notion of an economically bubbly market is numéraire-independent. However, we should point out that this fundamental value still depends on the class $\mathcal{U}$ of undefaultable strategies that we use in Definition 2.16.

2) Anticipating a concept from Section 4.2, $\pi^i_\tau$ is actually the superreplication price $\Pi_\tau(V_\tau(e^i) | \mathcal{U})$ at time $\tau$ of the contingent claim $V_\tau(e^i)$; see Definition 4.7 below.

Let us now show that our two notions of “bubbliness” coincide. The proof anticipates some results given below in Section 4.2 and can be skipped on a first reading. In particular, we need and use in the proof the identity (4.4).

Proposition 2.19. Suppose that $\mathcal{S}$ is dynamically viable. Then $\mathcal{S}$ is bubbly (in the sense of Definition 2.14) if and only if it is economically bubbly.

Proof. If $\mathcal{S}$ is economically bubbly, there are an asset $i \in \{1, \ldots, N\}$ and $\tau \in \mathcal{T}_{[0,T]}$ with $P[\pi^i_\tau(S) < S^i_\tau] = P[\pi^i_\tau(S) < S^i_\tau] > 0$. So $C := V_\tau(e^i) - \Pi_\tau(V_\tau(e^i) | \mathcal{U})$ is a nonzero contingent claim at time $\tau$, and $F(\cdot) := C(S^{(\delta)})V_\tau(\eta^S)(\cdot)$ is by (4.4) a nonzero contingent claim at time $T$. Lemma 4.9 below for $V_\tau(e^i)$ (at time $T$) and $V_\tau(\eta^S)$ (at time $\tau \leq T$) thus yields for $\delta > 0$ a strategy $\delta^\delta \in \mathcal{U}$ with $V_\tau(\delta^\delta) \geq V_\tau(e^i)$ and, by the definition of $C$ and (4.4), $V_\tau(\delta^\delta)(\cdot) \leq \Pi_\tau(V_\tau(e^i) | \mathcal{U})(\cdot) + \delta V_\tau(\eta^S)(\cdot) = V_\tau(e^i)(\cdot) - C(S^{(\eta^S)})V_\tau(\eta^S)(\cdot) + \delta V_\tau(\eta^S)(\cdot)$. Setting $\delta^\delta := \delta^\delta + C(S^{(\eta^S)})\eta^S1_{[\tau,T]}$ gives a strategy which is also in $\mathcal{U}$ because $C \geq 0$, and by construction, it satisfies $V_\tau(\delta^\delta) = V_\tau(\delta^\delta) + C(S^{(\eta^S)})V_\tau(\eta^S) \leq V_\tau(e^i) + \delta V_\tau(\eta^S)$ and $V_\tau(\delta^\delta) = V_\tau(\delta^\delta) + C(S^{(\eta^S)})V_\tau(\eta^S) \geq V_\tau(e^i) + F$, as in (2.8). This clearly shows that $e^i$ is not strongly maximal for $\mathcal{U}$, so that $\mathcal{S}$ is not dynamically efficient, hence bubbly.

Conversely, if $\mathcal{S}$ is bubbly (in the sense of Definition 2.14), it fails to be dynamically efficient, which by Corollary 5.3 below implies that there is an index $i \in \{1, \ldots, N\}$ such that $e^i$ fails to be strongly maximal for $\mathcal{U}_0$. Seeking a contradiction, suppose that $^*S^i_\tau = S^i_\tau$ $P$-a.s. for each $\tau \in \mathcal{T}_{[0,T]}$. Because $\mathcal{S}$ is dynamically viable, [23, Theorem 4.19] gives the existence of a strongly maximal strategy $\delta^* \in \mathcal{U}_0$ with $^*S^i_0 = V_0(\delta^*)(S) = V_0(e^i)(S) = S^i_0$.
and \( V_T(\vartheta^*) (S) \geq V_T(\vartheta^*)(S) = S_T^\vartheta \) \( \mathbb{P}\)-a.s. Hence \( V_T(\vartheta^*) (S) \geq S_T^\vartheta \) \( \mathbb{P}\)-a.s. for all \( \tau \in \mathcal{T}_{[0,T]} \) by the definition of superreplication prices, it follows that \( V_T(\vartheta^*) (S) - V_T(\vartheta^*)(S) \geq S_T^\vartheta - S_T^\vartheta = 0 \) \( \mathbb{P}\)-a.s. for each \( \tau \in \mathcal{T}_{[0,T]} \), and so \( \vartheta^* - \vartheta^1 \in \mathcal{U}_0 \). As \( V_0(\vartheta^* - \vartheta^1)(S) = 0 \), dynamic viability of \( \mathcal{S} \) directly yields \( V_T(\vartheta^*)(S) = 0 \) \( \mathbb{P}\)-a.s. for each \( \tau \in \mathcal{T}_{[0,T]} \). Thus \( V(\vartheta^*)(S) = V(\vartheta^1)(S) \) \( \mathbb{P}\)-a.s., and so \( \vartheta^1 \) is like \( \vartheta^* \) strongly maximal for \( \mathcal{U}_0 \). This is a contradiction. 

**Remark 2.20.** If \( \mathcal{S} \) fails to be dynamically viable, the nontrivial (“only if”) direction of Proposition 2.19 is wrong; this is not surprising because the bubble can come from an arbitrage in the market. For a concrete example, let \( \mathcal{S} \) be the market generated by \( S = (S^1, S^2) = (1, X_k)_{k \in \{0,1\}} \), where \( X_0 = 1 \) and \( X_1 \) takes the values 1 and 2 both with positive probability. Then \( \mathcal{S} \) clearly fails to be dynamically viable and hence is bubbly in the sense of Definition 2.14. However, it is straightforward to check that \( \ast S^1 = S^1 \) \( \mathbb{P}\)-a.s. and \( \ast S^2 = S^2 \) \( \mathbb{P}\)-a.s.

### 3 First examples

This section gives a number of examples to illustrate the concepts introduced so far, focussing on the (sometimes subtle) differences between different notions. Except for Example 3.3, we do not yet need here the dual characterisations from Section 5 below.

We start with an explicit example to show that our approach is more general than the classic setup of mathematical finance discussed in Example 2.2.

**Example 3.1 (A market that does not fit into the classic setup).** For \( i = 1, 2 \), take independent Brownian motions \( W^i = (W^i_t)_{t \in [0,T]} \) and stopping times \( \tau_i \) for the usual (augmented) filtration generated by \( (W^1, W^2) \), with \( \mathbb{P} \{ 0 < \tau_i < T \} > 0 \). Moreover, let \( X_i \) be random variables which are independent of \( (W^1, W^2) \) and satisfy \( \mathbb{P}[X_i = -1] = p_i \), \( \mathbb{P}[X_i = \alpha_i] = 1 - p_i \) with \( \alpha_i > 0 \), \( p_i \in (0,1) \), \( i = 1, 2 \), and \( \mathbb{P}[X_1 = -1, X_2 = -1] = 0 \). Define the one-jump processes \( N^i = (N^i_t)_{t \in [0,T]} \) by \( N^i_t = X_i \mathbb{I}_{\{t \geq \tau_i\}} \), \( i = 1, 2 \), and let \( (\mathcal{F}_t)_{t \in [0,T]} \) be the (augmented) filtration generated by \( (W^1, W^2, N^1, N^2) \). Let \( \mu_1, \mu_2 \in \mathbb{R} \) and \( \sigma_1, \sigma_2 > 0 \), and define the process \( S = (S^1_t, S^2_t)_{t \in [0,T]} \) by the SDEs

\[
\begin{align*}
\text{d}S^1_t &= S^1_{t-} \left( \mu_1 \text{ d}t + \sigma_1 \text{ d}W^1_t + \text{ d}N^1_t \right), \\
S^0_0 &= s_i > 0, i = 1, 2.
\end{align*}
\]

It is clear that for \( i = 1, 2 \), prior to \( \tau_i \), \( S^i \) is a geometric Brownian motion with drift \( \mu_i \) and volatility \( \sigma_i \). At \( \tau_i \), it either jumps to zero (if \( X_i = -1 \)) and stays there, or it jumps to \( (1 + \alpha_i)S^i_{\tau_i} \) (if \( X_i = \alpha_i \)) and evolves from there as a geometric Brownian motion with the same parameters \( \mu_i, \sigma_i \) as before the jump.

Note that both \( S^1 \) and \( S^2 \) may jump to zero with positive probability; but because of \( \mathbb{P}[X_1 = -1, X_2 = -1] = 0 \), at least one of them stays strictly positive until time \( T \). So (2.1) is satisfied for \( S \), and (2.3) also holds (with \( D = 1 \)) since \( S \) is a semimartingale. However, this example cannot be treated in the classic setup since no asset price process is guaranteed to remain positive with probability 1.

Here is an intuitive (but a bit extreme) situation where this model is natural. Suppose \( \tau_1 = \tau_2 =: \tau \) and \( X_1 = -X_2 \), which forces \( \alpha_1 = \alpha_2 = 1 \). Then both \( S^1 \) and \( S^2 \) evolve as (independent) geometric Brownian motions up to \( \tau \), which is interpreted as the random
and a fortiori strongly) maximal for any important announcement is made. One asset then drops to 0; the other instantaneously doubles its price and then continues as a GBM. Which of the assets “defaults” is determined by the value of $X_1$, which can be interpreted as a signal to the market at time $\tau$. For instance, $S^1, S^2$ could be the values of two firms competing for a monopoly in the same market sector, and $X_1 = 1$ means that firm 1 gets control of that sector—for example because it obtains the single licence available for a telecom market.

**Remark 3.2.** The above example has processes that jump to zero when they default. One could construct analogous examples with continuous processes that creep down to zero. Our reason for using a jump process is just that this gives a simpler construction.

Because of its purpose, Example 3.1 cannot be formulated in the classic setup. In contrast, all subsequent examples in this section have one asset whose price is identically 1; so they do not hinge on the fact that we work with a more general setting.

Our second example, essentially due to Delbaen/Schachermayer [5], exhibits a model where $\tau$ is determined by the value of $X_1$, which can be interpreted as a signal to the market at time an important announcement is made. One asset then drops to 0; the other instantaneously doubles its price and then continues as a GBM. Which of the assets “defaults” is determined by the value of $X_1$, which can be interpreted as a signal to the market at time $\tau$. For instance, $S^1, S^2$ could be the values of two firms competing for a monopoly in the same market sector, and $X_1 = 1$ means that firm 1 gets control of that sector—for example because it obtains the single licence available for a telecom market.

**Example 3.3 (A strategy which is not maximal).** Let $W = (W_t)_{t \in [0, T]}$ be a Brownian motion with respect to a given filtration $(\mathcal{F}_t)_{t \in [0, T]}$; this need not be generated by $W$. Consider the market $\mathcal{S}$ generated by $S = (S^1, S^2) = (1, X_t)_{t \in [0, T]}$, where $X$ is a three-dimensional Bessel process $BES^3$, i.e., the unique strong solution of the SDE

$$ dX_t = \frac{dt}{X_t} + dW_t, \quad X_0 = s_0 > 0. $$

We claim that the invest-and-keep strategy $e^1 = (1, 0)$ of the first asset fails to be (weakly and a fortiori strongly) maximal for $\mathcal{U}_0$; in other words, it is not a good idea in this market to put money into the bank account.

To see this, we first look at the invest-and-keep strategy $e^2 = (0, 1)$ of the second asset and note that $S^{(e^2)} = (1/X, 1)$. Itô’s formula shows that $Y := 1/X$ satisfies the SDE

$$ dY_t = -|Y_t|^2 dW_t, \quad Y_0 = 1/s_0. \tag{3.1} $$

It is well known that $Y$ is a local $\mathbb{P}$-martingale, and even a strict local $\mathbb{P}$-martingale. Due to Theorem 4.3 below, the market $\mathcal{S}$ is dynamically viable, and by Lemma 5.1 below, weak and strong maximality for each $\mathcal{U}_0$ are therefore equivalent.

To establish that $e^1$ is not maximal for $\mathcal{U}_0$, we use Theorem 4.12 in Herdegen [23]. As $S^{(e^2)} = (Y, 1)$ is a local $\mathbb{P}$-martingale and $V(e^2)(S^{(e^2)}) \equiv 1$ is a (true) $\mathbb{P}$-martingale, the numéraire strategy $e^2$ is maximal for $\mathcal{U}_0$ by [23, Theorem 4.12 (d)]. By [23, Theorem 4.12 (c)], it thus suffices to show that for every $Q \approx \mathbb{P}$ on $\mathcal{F}_T$ for which $S^{(e^2)}$ is a local $\mathbb{Q}$-martingale, $V(e^1)(S^{(e^2)}) = Y$ is a strict local $\mathbb{Q}$-martingale. But $Y$ satisfies the SDE (3.1), written for clarity as $dY_t = -|Y_t|^2 dW^P_t$, so that $(Y)_t = \int_0^t |Y_s|^2 ds$, and since $Y = 1/X$ is positive, we can write $W^P = \int -\frac{1}{|Y|^2} dY$. This is like $Y$ a continuous local $\mathbb{Q}$-martingale, and it has quadratic variation $(W^P)_t = \int_0^t \frac{1}{|Y|^2} d(Y)_s = t$, $t \in [0, T]$. Hence $W^P$ is also a $\mathbb{Q}$-Brownian motion $W^Q$, and so $Y$ has the same distribution under $Q$ as under $\mathbb{P}$. In particular, $Y$ is also a strict local $\mathbb{Q}$-martingale, and so $e^1$ fails to be maximal for $\mathcal{U}_0$. 

14
Of course, this example is classic; it is well known that the $BES^3$ process can be used to construct counterexamples, and the non-maximality of $e^1$ is also already pointed out in [6, Remark after Corollary 5] (even if maximality there has a slightly different meaning, as discussed in Remark 2.11). One small but neat novelty of the present example is that we do not need to assume that the filtration is generated by $W$ or by $X$.

**Remark 3.4.** Although the filtration in Example 3.3 is general, the model is quite special because it is complete for its own filtration. More precisely, the local $Q$-martingale $Y = 1/X$ has the predictable representation property in the (augmented) filtration $(\mathcal{F}^S_t)$ generated by $S$ (or, equivalently due to (3.1), by $W$ or by $Y$). Example 6.3 below exhibits another natural example which is genuinely incomplete.

We next turn to the difference between weak and strong maximality. Explicit and detailed examples of strategies which are weakly but not strongly maximal require quite a bit of work. One is given by Delbaen/Schachermayer [9, Example 9.7.7] or [4, Example 7.7]; another can be found in Hulley [27, Example 1.37]. We give below an even more extreme example, where every $\vartheta \in \mathcal{U}_0$ is weakly maximal for $\mathcal{U}_0$, but no $\vartheta \in \mathcal{U}_0$ is strongly maximal for $\mathcal{U}_0$. In view of Lemma 5.1 below, this market cannot be dynamically viable.

**Example 3.5 (Strategies which are weakly, but not strongly maximal).** This example is from Herdegen [23, Example 3.11]; it is also discussed in [24, Section 5.1]. We do not give full details of all the required computations, but merely outline the ideas and proofs.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space supporting a random variable $U$ that is uniformly distributed on $(0, 1)$. Set $\mathcal{F}^0 := \{U \leq u\} : u \in t)$ and $\mathcal{F}_t := \mathcal{F}^0_t \vee \sigma(\mathcal{N})$ for $t \in [0, 1]$, where $\mathcal{N}$ denotes the family of $\mathbb{P}$-nullsets in $\mathcal{F}^0$. Then consider the market $S$ generated by $S = (S^1, S^2) = (1, S^2_t)_{t \in [0,T]}$, where $S^2$ is a single-jump process explicitly given by $S^2_t := F_1(t) \mathbf{1}_{t < U} + F_2(U) \mathbf{1}_{t \geq U}$ and $F_1, F_2 : [0, 1] \to \mathbb{R}$ are the functions

$$F_1(t) := \begin{cases} 1 + s - \frac{2}{3}s^2, & t \in [0, 1/2] \\ \frac{4}{3} = F_1(1/2), & t \in (1/2, 1] \end{cases} \quad \text{and} \quad F_2(t) := \begin{cases} \frac{1}{3} + 3s - 2s^2, & t \in [0, 1/2] \\ \frac{4}{3} = F_2(1/2), & t \in (1/2, 1]. \end{cases}$$

So the process $S^2$ behaves almost deterministically—there is one single jump at a random time $U$ between 0 and 1, and at that time, the process jumps from one continuous trajectory (namely $F_1$) to another (namely $F_2$) and stays at the level $F_2(U)$.

In this model, each $\vartheta \in \mathcal{U}_0$ is weakly but not strongly maximal for $\mathcal{U}_0$. We only present a heuristic argument adapted from [24, Section 5.1] and refer to [23] for a formal proof.

First, $S^2$ is increasing until $U \wedge \frac{1}{2}$ with a slope strictly larger than $\frac{1}{2}$, and it has at $U$ a nonpositive jump, whose size is close to 0 for $U$ close to $\frac{1}{2}$ and equals 0 for $U \geq \frac{1}{2}$, which happens with probability $\frac{1}{2}$. The intuitive reason why 0 (and hence any other $\vartheta \in \mathcal{U}_0$) is not strongly maximal for $\mathcal{U}_0$ is that for each initial wealth $\varepsilon > 0$, one can go long in $S^2$ shortly prior to $\frac{1}{2}$ and buy as many shares as the undefaultability condition allows. If $U \geq \frac{1}{2}$, i.e., with probability $\frac{1}{2}$, one takes a profit. The other one starts trading (always strictly prior to $\frac{1}{2}$), the smaller the size of the jump, and the more shares one can buy. Thus, for each initial wealth $\varepsilon$, one can always get a final value of at least 1 on $\{U \geq \frac{1}{2}\}$.

The intuitive reason that each $\vartheta \in \mathcal{U}_0$ is nevertheless weakly maximal for $\mathcal{U}_0$ is a bit more subtle. We just explain the heuristic for the zero strategy. Improving 0 by a long
position in $S^2$ is not possible, since the downward jump in $S^2$ can occur an arbitrarily short time after one makes the investment, and then one ends up with a negative value, violating the nonnegativity condition from $\mathcal{U}_0$. On the other hand, improving $0$ by a short position in $S^2$ is also not possible, since with probability $\frac{1}{2}$, no jump occurs and one ends up again with a negative value.

For both viability and efficiency, we have introduced a static and a dynamic version. The next, very simple example shows that the two concepts are different, and also clarifies where the difference comes from.

**Example 3.6 (Static versus dynamic viability/efficiency).** Let $\mathcal{S}$ be the market generated by $S = (S^1, S^2) = (1, X_k)_{k \in \{0,1,2\}}$ with $X$ given by the following event tree, where each branch is assumed to have positive probability.

$$X : 1 \quad \frac{1}{2} \quad 2 \quad \frac{5}{2} \quad 3 \quad \frac{3}{2} \quad 1 \quad \frac{1}{2}$$

The key feature of this model is that from time 0 directly to the final time 2 and from every node at time 1 to time 2, the asset price $S^2 = X$ can go both strictly up and down. Hence each such “one-step” model is arbitrage-free, invest-and-keep trading (which precisely corresponds to trading in such a one-step model) cannot improve any given invest-and-keep strategy, and so $\mathcal{S}$ is statically efficient, and a fortiori statically viable. Note that this is because our invest-and-keep strategies must always be kept until the end ($T = 2$ here). However, from time 0 to time 1, $X$ stays the same or goes up; so the one-step model in this sub-tree already admits arbitrage, then so does the whole model $\mathcal{S}$, and so $\mathcal{S}$ is not dynamically viable (and a fortiori not dynamically efficient either). So this is a bubbly market, but not an interesting one, and both for viability and efficiency, the static version of the concept is weaker than the dynamic one.

**Example 3.7 (Bubble birth).** For incomplete markets, there can be situations where we see no bubble at time 0, but there is a bubble at some later time. This is called “bubble birth” by some authors (e.g. [30, 31, 1]), and it can come up in our framework in a very natural way. In abstract terms, we need a market generated by a process $S$ which is a true martingale prior to a suitable stopping time $\tau$, and a strict local martingale after $\tau$. To construct an explicit example and show rigorously that it has the desired features, we work with the class of single-jump local martingales introduced and studied in [24, 25]. This yields an intuitive description, allows concrete computations, and gives at the same time a precise control over sufficiently many equivalent local martingale measures (ELMMs).

On $(\Omega, \mathcal{F}, \mathbb{P})$, take a Brownian motion $W$ and an independent random variable $\gamma$ with values in $(0,1]$ and $0 < \mathbb{P}[\gamma = 1] < 1$; so the distribution of $\gamma$ has an atom at 1. Define filtrations $\mathcal{F}^W, \mathcal{F}^\gamma, \mathcal{F}$ for $0 \leq t \leq 1$ by $\mathcal{F}^W_t := \sigma(W_s : 0 \leq s \leq t), \mathcal{F}^\gamma_t = \sigma(\mathbb{1}_{\{\gamma \leq s\}} : 0 \leq s \leq t)$.
and \( \mathcal{F}_t = \mathcal{F}_t^W \cup \mathcal{F}_t^\gamma \cup \mathcal{N} \), where \( \mathcal{N} \) denotes the \( \mathbb{P} \)-nullsets in \( \mathcal{F}_t^W \cup \mathcal{F}_t^\gamma \). The market \( S \) is generated by \( S = (1, X) \), where \( X = (X_t)_{0 \leq t \leq 1} \) is the unique strong solution to the SDE
\[
dX_t = X_t \left( \mu \, dt + \sigma(t, \gamma) \, dW_t \right), \quad X_0 = 1,
\]
with \( \mu \in \mathbb{R} \) and \( \sigma : [0, 1]^2 \to [\sigma_0, \infty) \) for some \( \sigma_0 > 0 \) given by
\[
\sigma(t, v) = \sigma_0 \left( 1 + \frac{1}{1-t} \mathbb{1}_{\{v \leq t<1\}} \right).
\]
The random time \( \gamma \) is a stopping time in \( \mathcal{F}^\gamma \subseteq \mathcal{F} \), and \( X \) is before \( \gamma \) just a geometric Brownian motion. At time \( \gamma \), there is a jump in the volatility which then blows up until time 1 in such a way that \( X \) converges to 0. Intuitively, \( \gamma \) can be interpreted as the time when “the bubble is born”; see below for a more precise discussion.

As in Definition 2.17, we denote by \( ^*S_t \) the fundamental value or superreplication price of \( S \) at time \( t \). We claim that this is given, for \( t < 1 \), by \( ^*S_t = (1, ^*X_t) \) with
\[
^*X_t = X_t \mathbb{1}_{\{\gamma > t\}}. \quad (3.2)
\]
For “\( \leq \)” in (3.2), we note that \( X_1 = 0 \) on \( \{ \gamma \leq t \} \). The strategy \( \vartheta := \mathbb{1}_{[t, 1]} \mathbb{1}_{\{\gamma > t\}} \epsilon^2 \) thus has \( V_1(\vartheta)(S) = \mathbb{1}_{\{\gamma > t\}} X_1 = X_1 \) and \( V_1(\vartheta)(S) = \mathbb{1}_{\{\gamma > t\}} X_t \), and so we get \( ^*X_t \leq \mathbb{1}_{\{\gamma > t\}} X_t \).

To argue “\( \geq \)” in (3.2), we use the hedging duality from Kramkov [36]; this says that
\[
^*S_t = \text{ess sup} \{ \mathbb{E}_Q[S_1 | \mathcal{F}_t] : \text{Q is an ELMM for } S \}
= \left( 1, \text{ess sup} \{ \mathbb{E}_Q[X_1 | \mathcal{F}_t] : \text{Q is an ELMM for } X \} \right). \quad (3.3)
\]
(Note that 1 and \( X_1 \) are nonnegative; so the self-financing strategies resulting from the optional decomposition theorem in [36] are actually undefaultable and not only \( a \)-admissible for some \( a > 0 \).) Hence it is enough to exhibit for each \( \epsilon > 0 \) an ELMM \( \text{Q} \) for \( X \) with
\[
\mathbb{E}_Q[X_1 | \mathcal{F}_t] \geq (1 - \epsilon) X_t \mathbb{1}_{\{\gamma > t\}}. \quad (3.4)
\]
Define the local \( (\mathbb{P}, \mathcal{F}^W) \)-martingale \( Z^1 = (Z^1_s)_{0 \leq s \leq 1} \) by
\[
dZ^1_s = -Z^1_s \frac{\mu}{\sigma(s, \gamma)} \, dW_s, \quad Z^1_0 = 1.
\]
As \( \frac{\left| \frac{\mu}{\sigma(s, \gamma)} \right|}{\frac{\mu}{\sigma_0}} \leq 1 \), \( Z^1 \) satisfies Novikov’s condition and hence is a true \( (\mathbb{P}, \mathcal{F}^W) \)-martingale on \([0, 1]\). The change of measure corresponding to \( Z^1 \) eliminates the drift from \( X \) and turns \( X \) into a local martingale. Next, define the local \( (\mathbb{P}, \mathcal{F}^\gamma) \)-martingale \( Z^2 = (Z^2_s)_{0 \leq s \leq 1} \) by
\[
Z^2_s = \left( \frac{1 - \epsilon}{\mathbb{P}[\gamma > s]} + \frac{\epsilon}{\mathbb{P}[\gamma < 1]} \frac{\mathbb{P}[s < \gamma < 1]}{\mathbb{P}[\gamma > s]} \right) \mathbb{1}_{\{s< \gamma\}}
+ \frac{\epsilon}{\mathbb{P}[\gamma < 1]} \mathbb{1}_{\{s \leq \gamma, \gamma < 1\}} + \frac{1 - \epsilon}{\mathbb{P}[\gamma = 1]} \mathbb{1}_{\{s = 1 = \gamma\}}; \quad (3.5)
\]
the corresponding change of measure changes the distribution of \( \gamma \), but leaves \( W \) unchanged and hence keeps \( X \) a local martingale. For the true martingale property of \( Z^2 \), note that \( Z^2 = \mathcal{M}^G F \) in the notation of [24], where \( F(t) = \frac{1 - \epsilon}{1-G(t)} + \frac{\epsilon}{G(1-\epsilon)-G(t)} \) and
$G$ is the distribution function of $\gamma$. Now $F$ is clearly (locally) absolutely continuous with respect to $G$, we have $\Delta G(1) = P[\gamma = 1] > 0$ by the assumption on $\gamma$, and

$$M_1^G F = Z_1^2 = \frac{\epsilon}{P[\gamma < 1]} \mathbb{I}_{\{\gamma < 1\}} + \frac{1 - \epsilon}{P[\gamma = 1]} \mathbb{I}_{\{\gamma = 1\}} \quad (3.6)$$

from (3.5) is bounded, hence integrable. Therefore it follows from [24, Theorem 3.5 (c)] that $Z^2 = M^G F$ is a true $(P, F^\gamma)$-martingale on $[0, 1]$.

Define $Q \approx P$ by $dQ = Z_2^2 Z_1^1 dP$. This will be an ELMM for $X$ in $F$ if we can check that $Z^2 Z^1$ is a (true) $(P, F)$-martingale and $Z^2 Z^1 X$ is a local $(P, F)$-martingale. The first claim follows from [25, Lemma A.1 (a)] with $\epsilon < 1$ on the second claim, we use that as result also holds for a completed filtration and for a general distribution for $\gamma$. For the second claim, we use that as $Z^2 Z^1 X$ is a continuous (hence special) $(P, F)$-semimartingale on $[0, 1]$, it suffices to show that it is a (local) $(P, F)$-martingale on $[0, u]$ for each fixed $u < 1$. But $Z^2 Z^1 X$ is even a true $(P, F)$-martingale on $[0, u]$ by [25, Lemma A.1 (a)] applied on $[0, u]$, with $Y^1 = Z^1 X$, and $Y^2 = Z^2$ there.

Now fix $t < 1$. On the set $\{\gamma > t\}$, we have $Z_1^1 X_t \mathbb{I}_{\{\gamma > t\}} = \mathcal{E}(\sigma_0 - \frac{\mu}{\sigma_0} W) t \mathbb{I}_{\{\gamma > t\}}$ because $\sigma(s, \gamma) \equiv \sigma_0$ up to time $t$. In the same way, $Z_1^1 X_t \mathbb{I}_{\{\gamma = 1\}} = \mathcal{E}(\sigma_0 - \frac{\mu}{\sigma_0} W) t \mathbb{I}_{\{\gamma = 1\}}$, and $X_t = 0$ on $\{\gamma < 1\}$. Combining this with (3.6) and $\mathbb{I}_{\{\gamma = 1\}} \mathbb{I}_{\{\gamma > t\}}$ yields

$$Z_t^2 Z_t^1 X_t = \frac{1 - \epsilon}{P[\gamma = 1]} \mathbb{I}_{\{\gamma = 1\}} \mathbb{I}_{\{\gamma > t\}} Z_t^1 X_t \mathcal{E}(\sigma_0 - \frac{\mu}{\sigma_0} W) t \mathcal{E}(\sigma_0 - \frac{\mu}{\sigma_0} W) t,$$

and the ratio is independent from $\gamma$. The Bayes rule and $F_t = F_t^{\gamma} \vee F_t^W$ $P$-a.s. then give

$$E_Q[X_t | F_t] = \frac{E[Z_t^2 Z_t^1 X_t | F_t]}{Z_t^2 Z_t^1} = \frac{1}{Z_t^2 Z_t^1} \frac{1 - \epsilon}{P[\gamma = 1]} Z_t^1 X_t \mathbb{I}_{\{\gamma > t\}} \mathbb{I}_{\{\gamma = 1\}} \mathcal{E}(\sigma_0 - \frac{\mu}{\sigma_0} W) t \mathcal{E}(\sigma_0 - \frac{\mu}{\sigma_0} W) t | F_t] = \frac{1}{Z_t^2 Z_t^1} \frac{1 - \epsilon}{P[\gamma = 1]} X_t \mathbb{I}_{\{\gamma > t\}} P[\gamma = 1 \mid F_t^\gamma] \mathcal{E}(\sigma_0 - \frac{\mu}{\sigma_0} W) t \mathcal{E}(\sigma_0 - \frac{\mu}{\sigma_0} W) t | F_t^W] = \frac{1}{Z_t^2 Z_t^1} \frac{1 - \epsilon}{P[\gamma = 1]} X_t \mathbb{I}_{\{\gamma > t\}} P[\gamma = 1 \mid F_t^\gamma] P[\gamma > t] = \frac{1}{Z_t^2 Z_t^1} \frac{1 - \epsilon}{P[\gamma = 1]} X_t \mathbb{I}_{\{\gamma > t\}} P[\gamma > t], \quad (3.7)$$

because $\{\gamma > t\}$ is an atom of $F_t^\gamma$ and the stochastic exponential is a $(P, F^W)$-martingale. On the other hand, we have from (3.5)

$$Z_t^2 = \frac{1}{P[\gamma > t]} \mathbb{I}_{\{\gamma > t\}} \left(1 - \epsilon + \epsilon \frac{P[t < \gamma < 1]}{P[\gamma < 1]}\right) \leq \frac{1}{P[\gamma > t]} \mathbb{I}_{\{\gamma > t\}}.$$

Plugging this into (3.7) yields (3.4) and completes the proof of (3.2).

For any $t < 1$ with $P[\gamma > t] < 1$, the relation $^*X_t = X_t \mathbb{I}_{\{\gamma > t\}}$ in (3.2) now implies that $P[^*X_t < X_t] > 0$ so that we see a bubble at time $t$. However, $^*X_0 = X_0$ since $\gamma > 0$; so we see no bubble at time 0. Moreover, $X$ is strictly positive on $[0, 1]$ so that we also get

$$\gamma = \inf\{t \in [0, 1] : ^*S_t \neq S_t\}.$$
This shows that \( \gamma \) is indeed the time when “the bubble is born”. Note that the property \( P\left[ X_t < X_t \right] > 0 \) together with (3.3) shows that \( S \) is on the interval \([0, 1]\) a strict local \( Q \)-martingale under every \( \text{ELMM} \) \( Q \) for \( S \), although we have \( \sup \{ E_Q[S_t] : \text{Q ELMM for } S \} = S_0 \). Thus the strict local martingale property only appears at strictly positive times; one cannot detect it at time 0.

4 Absence of arbitrage and superreplication prices

In this section, we connect viability to absence of arbitrage and to a dual description which involves martingale properties. We then recall from [23] the concept and some properties of contingent claims in our numéraire-independent framework.

4.1 Viability and absence of arbitrage

We first show that many concepts and results simplify in finite discrete time. This is no surprise since that setting is well known to be much easier than a model with infinitely many trading dates.

**Definition 4.1.** We say that the market \( S \) satisfies no arbitrage (NA) if there is no strategy \( \vartheta \in U_0(S) \) satisfying, for some (or equivalently all) \( S \in S \),

\[
V_0(\vartheta)(S) = 0, \quad V_T(\vartheta)(S) \geq 0 \text{ P-a.s. and } P[V_T(\vartheta)(S) > 0] > 0. \tag{4.1}
\]

For finite discrete time and the classic setup as in Example 2.2, the above is the standard classic definition of absence of arbitrage; see for example Elliott/Kopp [12, Definition 2.2.3 and the subsequent section]. For a continuous-time model in the classic setup, the above condition was studied under the name (NA+) by Hulley [27]. Note also that Definition 4.1 is “numéraire-independent”, and that the explicit requirement \( V_T(\vartheta)(S) \geq 0 \) P-a.s. is redundant since we have \( \vartheta \in U_0 \).

**Proposition 4.2.** Let \( S \) be a market in finite discrete time and recall the nondegeneracy assumption (2.1). Then the following are equivalent:

1) \( S \) satisfies NA.

2) \( S \) is dynamically viable.

3) \( S \) is dynamically efficient.

4) There exist a numéraire strategy \( \eta \) and a probability measure \( Q \approx P \) on \( \mathcal{F}_T \) such that the \( V(\eta) \)-discounted price process \( S^{(\eta)} = \frac{S}{V(\eta)(S)} \) is a \( Q \)-martingale.

5) For each numéraire strategy \( \eta \), there exists a probability measure \( Q \approx P \) on \( \mathcal{F}_T \) such that the \( V(\eta) \)-discounted price process \( S^{(\eta)} = \frac{S}{V(\eta)(S)} \) is a \( Q \)-martingale.

**Proof.** We show below in Theorem 5.2 that 4) implies 3), and it is clear that 5) implies 4) and that 3) implies 2). Next, 2) implies that 0 is weakly maximal for \( U_0 \), which is in
turn clearly equivalent to \( S \) satisfying NA, and so we obtain 1). So it only remains to argue that 1) implies 5), and this is where we exploit the setting of finite discrete time.

Let \( \eta \) be any numéraire strategy; by \((2.1)\), the market portfolio \( \eta^S = 1 \) is one example. Then we have \((4.1)\), with \( S \) replaced by \( X := S^{(0)} \). We claim that for \( t \in \{1, \ldots, T\} \), there can be no \( \mathcal{F}_{t-1} \)-measurable \( \mathbb{R}^N \)-valued random vector \( \xi \) such that

\[
\xi \cdot (X_t - X_{t-1}) \geq 0 \text{ P-a.s. and } \mathbb{P}[\xi \cdot (X_t - X_{t-1}) > 0] > 0.
\]

Indeed, if we have such \( t \) and \( \xi \), we can define an \( \mathbb{R}^N \)-valued predictable process \( \vartheta \) by

\[
\vartheta_k = \begin{cases} 
0, & k \leq t - 1, \\
\xi - (\xi \cdot X_{t-1})\eta_t, & k = t, \\
\xi \cdot (X_t - X_{t-1})\eta_k, & k > t.
\end{cases}
\]

(This is the usual strategy of investing \( \xi \) into the risky assets \( X \) from time \( t-1 \) to \( t \) and then putting the proceeds into the numéraire.) It is straightforward to check that \( \vartheta \) is in \( \mathcal{U}_0 \) due to \((4.2)\), and because \( V(\eta)(S^{(0)}) \equiv 1 \), \( \vartheta \) satisfies

\[
V_0(\vartheta)(S^{(0)}) = 0 \quad \text{and} \quad \mathbb{P}[V_T(\vartheta)(S^{(0)}) > 0] = \mathbb{P}[\xi \cdot (X_t - X_{t-1}) > 0] > 0,
\]

contradicting \((4.1)\). Thus, applying \cite[Proposition 5.11 and Theorem 5.16]{17} to the model \((1, X)\) gives \( Q \approx \mathbb{P} \) on \( \mathcal{F}_T \) such that \( X = S^{(0)} \) is a \( \mathbb{Q} \)-martingale, and we have 5). \( \square \)

Despite its simplicity, Proposition 4.2 already illustrates a key difference to the results in the classic setup of mathematical finance from Example 2.2. To see this, recall the well-known primal and dual objects in that setting. One starts with an \( \mathbb{R}^d \)-valued process \( X \) modelling the prices of \( d \) risky assets, expressed in units of a further asset labelled 0 and called bank account. Then \textit{primal objects} are self-financing strategies, and these can by Remark 2.5 be parametrised by pairs \((v_0, \psi)\) of initial wealths \( v_0 \in \mathbb{L}^0(\mathcal{F}_0) \) and \( \mathbb{R}^d \)-valued predictable \( X \)-integrable processes \( \psi \) describing the holdings in the risky assets. \textit{Dual objects} are then equivalent local martingale measures (ELMMs) for \( X \), i.e., probability measures \( Q \approx \mathbb{P} \) on \( \mathcal{F}_T \) such that \( X \) is a local \( \mathbb{Q} \)-martingale. (We do not need \( \sigma \)-martingales when price processes are nonnegative.) Finally, the fundamental theorem of asset pricing (FTAP) says that absence of arbitrage for \( X \) under \( \mathbb{P} \) (in the sense that \( X \) satisfies the condition NFLVR of no free lunch with vanishing risk) is equivalent to the existence of an ELMM \( Q \approx \mathbb{P} \) on \( \mathcal{F}_T \) for \( X \).

As one can see above, the very first step in the classic setup is to choose and fix a “numéraire”, namely asset 0 (the “bank account”). All subsequent definitions and results depend on this, and some concepts even cannot be defined without it. (A more thorough discussion can be found in the introduction of Herdegen \cite{23}.) In our more general setup, all \( N = d + 1 \) assets in \( S \) are treated symmetrically and we neither do, nor want to, choose a priori any particular numéraire. Primal objects are again self-financing strategies, now parametrised by \( \mathbb{R}^N \)-valued predictable \( S \)-integrable processes \( \vartheta \) which satisfy the self-financing constraint \((2.4)\). But dual objects, as already seen in \cite{23} and clearly illustrated by Proposition 4.2, are now \textit{pairs} \((\eta, Q)\), where \( \eta \) is a numéraire strategy and \( Q \) is an ELMM for the numéraire representative \( S^{(0)} \).

Now recall that \( S \) is nonnegative and describes a numéraire market due to \((2.1)\). Hence we have from Herdegen \cite{23} the following numéraire-independent version of the FTAP.
Theorem 4.3. The following are equivalent:

1) $S$ is dynamically viable.

2) $S$ satisfies numéraire-independent no-arbitrage (NINA), i.e., the zero strategy $0$ is strongly maximal for $U_0$.

3) There exists a pair $(\eta, Q)$, where $\eta$ is a numéraire strategy and $Q \approx P$ on $\mathcal{F}_T$, such that $S^{(n)}$ is a local $Q$-martingale.

4) There exist a representative $\bar{S} \in S$ and $Q \approx P$ on $\mathcal{F}_T$ such that $\bar{S}$ is a local $Q$-martingale.

Note that in general $S^{(n)}$ (or $\bar{S}$) may fail to be a true $Q$-martingale.

Proof. The equivalence of 2), 3) and 4) follows from the equivalence of (a), (b) and (d) in Herdegen [23, Theorem 4.9]. 1) implies 2) by the Definition 2.12 of dynamic viability, and that 2) implies 1) is shown below in Corollary 4.11. \qed

Remark 4.4. 1) As mentioned before, we get local martingales and not $\sigma$-martingales because our prices are nonnegative.

2) The FTAP in the classic setup $S = (1, X)$ with $N = d + 1$ (and $X \geq 0$, for easier comparison) states that $X$ satisfies NFLVR if and only if there exists $Q \approx P$ on $\mathcal{F}_T$ such that $X$ is a local $Q$-martingale. Note two crucial differences to Theorem 4.3: the numéraire $S^0$ is fixed a priori and, more importantly, even the formulation of the no-arbitrage condition NFLVR depends on the chosen currency unit, via the class of admissible strategies; see also the discussion in [23, Introduction]. There is no definition of NFLVR without first fixing the numéraire, and hence a result or formulation like Theorem 4.3 is impossible in the classic setup.

4.2 Contingent claims and superreplication prices

For providing dual characterisations of primal notions like efficiency and bubbles, so-called superreplication prices turn out to be very useful. They are also necessary for a valuation of financial contracts in a numéraire-independent way, but we do not address that aspect in the present paper. This section introduces or recalls some of the required concepts; more details and information can be found in [23, Section 2.5] or [21, Sections 4 and 5].

Definition 4.5. An improper contingent claim at time $\tau \in \mathcal{T}^{[0,T]}$ for the market $S$ is a map $F : S \rightarrow \mathbb{L}^0_+(\mathcal{F}_\tau)$ satisfying the exchange rate consistency condition

$$F(DS) = D_\tau F(S) \quad \text{P-a.s., for all } S \in S \text{ and all } D \in \mathcal{D}. \quad (4.3)$$

$F$ is called a contingent claim at time $\tau$ if it is valued in $\mathbb{L}^0_+(\mathcal{F}_\tau)$, and strictly positive if it is valued in $\mathbb{L}^{0+}_+(\mathcal{F}_\tau)$.

A contingent claim $F$ in our setup assigns to each representative $S \in S$ (which corresponds to a choice of currency unit) a payoff $F(S)$ at time $\tau$ (in the same unit), which is an $\mathcal{F}_\tau$-measurable random variable. The simplest example is the value $V_\tau(\vartheta)$ at time $\tau$ of any
self-financing strategy \( \vartheta \); (4.3) here follows from (2.5). For the canonical and most general example, we choose a pair \((g, S) \in \mathbf{L}_+^1(\mathcal{F}_\tau) \times \mathcal{S}\) and define \(F\) by \(F(S') = F(DS) := D_\tau g\) for any \(S' = DS\) in \(\mathcal{S}\); this clearly satisfies (4.3). Then \(g\) represents a payoff in the currency unit corresponding to \(S\), and we call \(F\) the contingent claim at time \(\tau\) induced by \(g\) with respect to \(S\). Like the self-financing condition in (2.6), we can extend (4.3) to arbitrary representatives \(\tilde{S} = S/\tilde{D}\) by setting \(F(\tilde{S}) := F(S)/\tilde{D}_\tau\).

**Remark 4.6.** It is important to distinguish clearly between a contingent claim \(F(\cdot)\) (in the above sense), which is a mapping with the property (4.3), and the corresponding payoff \(F(S)\) (in the units corresponding to the representative \(S\)), which is a random variable. In particular, the product of two contingent claims or a constant \(c \in \mathbb{R}\) are not contingent claims. However, for every numéraire strategy \(\eta\), we have the identity

\[ F(\cdot) = F(S^{(\eta)})V_\tau(\eta)(\cdot) \tag{4.4} \]

for every contingent claim \(F\) at time \(\tau\). Indeed, (4.4) holds for \(S := S^{(\eta)}\) due to (4.3) because \(V(\eta)(S^{(\eta)}) \equiv 1\) by (2.7), and then for general \(S\) due to (4.3) because \(V_\tau(\eta)\) is a contingent claim at time \(\tau\).

**Definition 4.7.** Let \(\sigma \leq \tau \in \mathcal{T}_{[0,T]}\) be stopping times, \(\Gamma\) a strategy cone on \([\sigma, T]\) and \(F\) a contingent claim at time \(\tau\). The superreplication price of \(F\) at time \(\sigma\) for \(\Gamma\) is the mapping \(\Pi_\sigma(F | \Gamma) : \mathcal{S} \to \mathbf{L}_+^0(\mathcal{F}_\sigma)\) defined by

\[ \Pi_\sigma(F | \Gamma)(S) := \inf \left\{ v \in \mathbf{L}_+^0(\mathcal{F}_\sigma) : \exists \vartheta \in \Gamma \text{ such that } \mathbb{P}\text{-a.s. on } \{ v < \infty \}, V_\tau(\vartheta)(S) \geq F(S) \text{ and } V_\sigma(\vartheta)(S) \leq v \right\}. \tag{4.5} \]

It is not difficult to check that \(\Pi_\sigma(F | \Gamma)\) is an improper contingent claim at time \(\sigma\). The following result lists some other basic properties. Note that these are properties of functions on \(\mathcal{S}\), and that they are all numéraire-independent in the (usual) sense that they hold for some \(S \in \mathcal{S}\) if and only if they hold for all \(S' \in \mathcal{S}\); this is due to the exchange rate consistency property (4.3). The proofs are straightforward and hence omitted.

**Proposition 4.8.** Let \(\sigma \leq \tau \in \mathcal{T}_{[0,T]}\) be stopping times, \(\Gamma\) a strategy cone on \([\sigma, T]\) and \(F, F_1, F_2, G\) contingent claims at time \(\tau\) with \(F \leq G\) \(\mathbb{P}\)-a.s. Let \(c_\sigma\) be a nonnegative \(\mathcal{F}_\sigma\)-measurable random variable. Then we have

\[
\Pi_\sigma(F | \Gamma) \leq \Pi_\sigma(G | \Gamma) \quad \text{(monotonicity),}
\]

\[
\Pi_\sigma(c_\sigma F | \Gamma) = c_\sigma \Pi_\sigma(F | \Gamma) \quad \text{(positive } \mathcal{F}_\sigma\text{-homogeneity),}
\]

\[
\Pi_\sigma(F_1 + F_2 | \Gamma) \leq \Pi_\sigma(F_1 | \Gamma) + \Pi_\sigma(F_2 | \Gamma) \quad \text{(subadditivity).}
\]

Note that positive \(\mathcal{F}_\sigma\)-homogeneity implies \(\Pi_\sigma(1_{A_\sigma} F | \Gamma) = 1_{A_\sigma} \Pi_\sigma(F | \Gamma)\) for \(A_\sigma \in \mathcal{F}_\sigma\). For conditional risk measures, this is called locality or the local property. However, in contrast to risk measures, we do not have \(\Pi_\sigma(F + C_\sigma | \Gamma) = C_\sigma + \Pi_\sigma(F | \Gamma)\) (the “cash-additivity analogue”) for contingent claims \(C_\sigma\) at time \(\sigma\).

The next auxiliary technical result can be used to approximate superreplication prices.
Lemma 4.9. Let \( \sigma \leq \tau \in \mathcal{T}_{[0,T]} \) be stopping times, \( \Gamma \) a strategy cone on \([\sigma,T]\) and \( F \) a contingent claim at time \( \tau \) with \( \Pi_\sigma(F|\Gamma) < \infty \) \( \mathbb{P} \)-a.s. (meaning that \( \Pi_\sigma(F|\Gamma)(S) < \infty \) \( \mathbb{P} \)-a.s. for some or, equivalently by (4.3), for all \( S \in \mathcal{S} \)). Then for all \( \delta > 0 \) and all strictly positive contingent claims \( C \) at time \( \sigma \), there exists a strategy \( \vartheta \in \Gamma \) satisfying

\[
V_\vartheta(\vartheta) \geq F \quad \mathbb{P}\text{-a.s.} \quad \text{and} \quad V_\vartheta(\vartheta) \leq \Pi_\sigma(F|\Gamma) + \delta C \quad \mathbb{P}\text{-a.s.}
\]

Proof. First, for any strictly positive contingent claim \( C \) at time \( \sigma \), there exists an \( S' \in \mathcal{S} \) with \( C(S') = 1 \) \( \mathbb{P} \)-a.s. Indeed, set \( D_T := 1/C(S) \in \mathbf{L}^0_+(\mathcal{F}_\sigma) \subseteq \mathbf{L}^0_{++}(\mathcal{F}_T) \) (for some fixed \( S \in \mathcal{S} \), take \( \mathbb{Q} \approx \mathbb{P} \) on \( \mathcal{F}_T \) with \( \mathbb{E}_\mathbb{Q}[D_T] < \infty \) and define \( D \in \mathcal{D} \) as an RCLL version of the \( \mathbb{Q} \)-martingale \( D_t = \mathbb{E}_\mathbb{Q}[D_T | \mathcal{F}_t], \ t \in [0,T] \). Note that \( D_\sigma = \mathbb{E}_\mathbb{Q}[D_T | \mathcal{F}_\sigma] = D_T \mathbb{P}\text{-a.s.} \) since \( \mathcal{F}_T \) is \( \mathcal{F}_\sigma \)-measurable. For \( S' := DS \), the exchange rate consistency (4.3) for \( C \) gives \( C(S') = C(DS) = D_\sigma C(S) = D_T/D_T = 1 \) \( \mathbb{P} \)-a.s.

Now take \( \delta > 0 \) and note that \( \Pi_\sigma(F|\Gamma) \) is a contingent claim at time \( \sigma \) and \( C(S') = 1 \). So by the exchange rate consistency (4.3), it suffices to show that there is \( \vartheta \in \Gamma \) with

\[
V_\vartheta(\vartheta)(S') \geq F(S') \quad \mathbb{P}\text{-a.s.} \quad \text{and} \quad V_\vartheta(\vartheta)(S') \leq \Pi_\sigma(F|\Gamma)(S') + \delta \quad \mathbb{P}\text{-a.s.} \quad (4.6)
\]

The set \( \mathcal{V} := \{ v \in \mathbf{L}^0_+(\mathcal{F}_\sigma) : \exists \vartheta \in \Gamma \text{ with } V_\vartheta(\vartheta)(S') \geq F(S') \text{ and } V_\vartheta(\vartheta)(S') \leq v, \mathbb{P} \text{-a.s.} \} \) is nonempty due to \( \Pi_\sigma(F|\Gamma) < \infty \), and also closed under taking minima. Indeed, if \( v_i \in \mathcal{V} \) and \( \vartheta_i \in \Gamma \) due to \( V_{\vartheta_i}(\vartheta_i)(S') \geq F(S') \) \( \mathbb{P} \)-a.s. and \( V_{\vartheta_i}(\vartheta_i)(S') \leq v_i \mathbb{P} \)-a.s. for \( i = 1,2 \), then \( \vartheta := \vartheta_1 \mathbb{1}_{\{v_1 \leq v_2\}} + \vartheta_2 \mathbb{1}_{\{v_1 > v_2\}} \in \Gamma \), and \( V_\vartheta(\vartheta)(S') \geq F(S') \) and \( V_\vartheta(\vartheta)(S') \leq v^1 \land v^2 \), \( \mathbb{P} \)-a.s. So there is a decreasing sequence \( (v_n)_{n \in \mathbb{N}} \) with \( \lim_{n \to \infty} v_n = \text{ess inf} \mathcal{V} = \Pi_\sigma(F|\Gamma)(S') \mathbb{P}\text{-a.s.} \), and \( B_n := \{ v_n \leq \Pi_\sigma(F|\Gamma)(S') + \delta \} \), \( B_0 := \emptyset \) and \( A_n := B_n \setminus B_{n-1} \) yields a partition \( (A_n)_{n \in \mathbb{N}} \) of \( \Omega \) into \( \mathcal{F}_\sigma \)-measurable sets. Take \( (\vartheta^n)_{n \in \mathbb{N}} \in \Gamma \) with \( V_{\vartheta^n}(\vartheta^n)(S') \geq F(S') \mathbb{P} \)-a.s. and \( V_{\vartheta^n}(\vartheta^n)(S') \leq v_n \mathbb{P} \)-a.s. for all \( n \). Then \( \vartheta := \sum_{n=1}^\infty \mathbb{1}_{A_n} \vartheta^n \) is in \( \Gamma \) and satisfies (4.6). \( \square \)

Recall that for \( \sigma \in \mathcal{T}_{[0,T]} \) and a strategy cone \( \Gamma \) on \([\sigma,T]\), a strategy \( \vartheta \in \Gamma \) is strongly maximal for \( \Gamma \) if there is no random variable \( f \in \mathbf{L}^0_+(\mathcal{F}_T) \setminus \{0\} \) such that for all \( \varepsilon > 0 \), there exists a strategy \( \vartheta \in \Gamma \) with \( V_T(\vartheta)(S) \geq V_T(\vartheta)(S) + f \mathbb{P} \)-a.s. and \( V_{\vartheta}(\vartheta)(S) \leq V_\vartheta(\vartheta)(S) + \varepsilon \mathbb{P} \)-a.s. This can now be reformulated more compactly: \( \vartheta \in \Gamma \) is strongly maximal for \( \Gamma \) if and only if there is no nonzero contingent claim \( F \) at time \( T \) such that

\[
\Pi_\sigma \left( V_T(\vartheta) + F \right| \Gamma \right)(S) \leq V_\vartheta(\vartheta)(S) \quad \mathbb{P}\text{-a.s.}
\]

This makes it clear and concise that for a maximal strategy, one cannot have more at the end without adding something at the beginning.

We next show that superreplication prices for undefaultable strategies are time-consistent, using that the family of all \( \mathcal{U}_\sigma := \mathcal{U}_\sigma(S) \) is itself time-consistent; see Section 2.1.

Proposition 4.10. Let \( \sigma_1 \leq \sigma_2 \leq \tau \in \mathcal{T}_{[0,T]} \) be stopping times and \( F \) a contingent claim at time \( \tau \) with \( \Pi_{\sigma_2}(F|\mathcal{U}_{\sigma_2}) < \infty \) \( \mathbb{P} \)-a.s. Then

\[
\Pi_{\sigma_1}(F|\mathcal{U}_{\sigma_1}) = \Pi_{\sigma_1} \left( \Pi_{\sigma_2}(F|\mathcal{U}_{\sigma_2}) \right| \mathcal{U}_{\sigma_1}) \quad \mathbb{P}\text{-a.s.} \quad (4.7)
\]

Proof. Denote the left- and right-hand sides of (4.7) by \( L \) and \( R \) respectively. For "\( \leq \)" it suffices to show the inequality for each (equivalently for some) \( S \in \mathcal{S} \) on the set

23
A := \{ R(S) < \infty \} \in \mathcal{F}_{\sigma_1} \subseteq \mathcal{F}_{\sigma_2}. \) By positive \( \mathcal{F}_{\sigma_i} \)-homogeneity, we may thus replace \( F \) by \( FI_{\{R<\infty\}} \), or equivalently assume without loss of generality that \( R < \infty \) \( \mathbb{P} \)-a.s. Analogously, for proving “≥”, we may assume that \( L < \infty \) \( \mathbb{P} \)-a.s.

“≤”: Fix a numéraire strategy \( \eta \). Take \( \delta > 0 \) and denote by \( C^i \) the contingent claim at time \( \sigma_i \) satisfying \( C^i(S^{(n)}) = 1, \ i = 1, 2 \); so \( C^i \) is induced at time \( \sigma_i \) by the constant 1 with respect to \( S^{(n)} \). By Lemma 4.9, there exist \( \vartheta^i \in \mathcal{U}_{\sigma_i}, i = 1, 2 \), satisfying

\[
V_{\sigma_2}(\vartheta^1) \geq \Pi_{\sigma_2}(F \mid \mathcal{U}_{\sigma_2}) \ \mathbb{P}\text{-a.s.} \quad \text{and} \quad V_{\sigma_1}(\vartheta^1) \leq R + \delta C^1 \ \mathbb{P}\text{-a.s.,} \tag{4.8}
\]

\[
V_{\sigma_2}(\vartheta^2) \geq F \ \mathbb{P}\text{-a.s.} \quad \text{and} \quad V_{\sigma_2}(\vartheta^2) \leq \Pi_{\sigma_2}(F \mid \mathcal{U}_{\sigma_2}) + \delta C^2 \ \mathbb{P}\text{-a.s.} \tag{4.9}
\]

By (2.7), the choice of \( C^2 \), (4.8) and (4.9),

\[
V_{\sigma_2}(\vartheta^1 + \delta \eta)(S^{(n)}) = V_{\sigma_2}(\vartheta^1)(S^{(n)}) + \delta C^2(S^{(n)}) \geq V_{\sigma_2}(\vartheta^2)(S^{(n)}) \ \mathbb{P}\text{-a.s.} \tag{4.10}
\]

Set \( \vartheta := (\vartheta^1 + \delta \eta)1_{[\sigma_1, \sigma_2]} + (\vartheta^2 + V_{\sigma_2}(\vartheta^1 + \delta \eta - \vartheta^2))(S^{(n)})\eta \chi_{[\sigma_2, T]} \). From (4.10) and the fact that \( \vartheta^i \in \mathcal{U}_{\sigma_i} \) and \( \eta \in \mathcal{U}_0 \subseteq \mathcal{U}_{\sigma_i} \), it is easy to check that \( \vartheta \in \mathcal{U}_{\sigma_1} \). Moreover, the definition of \( \vartheta \) gives by (4.9), (4.8) and (4.10) that \( V_{\sigma_2}(\vartheta) \geq F \ \mathbb{P}\text{-a.s.} \) and \( V_{\sigma_1}(\vartheta) \leq R + 2\delta C^1 \ \mathbb{P}\text{-a.s.} \). Thus \( \Pi_{\sigma_1}(F \mid \mathcal{U}_{\sigma_1}) \leq R + 2\delta C^1 \) by (4.5), and letting \( \delta \downarrow 0 \) yields the claim.

“≥”: Fix \( \delta > 0 \) and a strictly positive contingent claim \( C \) at time \( \sigma_1 \). By Lemma 4.9, there exists \( \vartheta \in \mathcal{U}_{\sigma_1} \) satisfying \( V_{\sigma_2}(\vartheta) \geq F \ \mathbb{P}\text{-a.s.} \) and \( V_{\sigma_1}(\vartheta) \leq L + \delta C \ \mathbb{P}\text{-a.s.} \). So the definition of superreplication prices gives first \( \Pi_{\sigma_2}(F \mid \mathcal{U}_{\sigma_2}) \leq \Pi_{\sigma_2}(\vartheta) \ \mathbb{P}\text{-a.s.} \) and then

\[
R = \Pi_{\sigma_1} \left( \Pi_{\sigma_2}(F \mid \mathcal{U}_{\sigma_2}) \bigg| \mathcal{U}_{\sigma_1} \right) \leq V_{\sigma_1}(\vartheta) \leq L + \delta C \ \mathbb{P}\text{-a.s.}
\]

The claim follows by letting \( \delta \downarrow 0 \).

Valuation by superreplication is well known to be consistent over time in the classic setup; see Föllmer/Schied [17, Example 11.2.4] for discrete or Klöppel/Schweizer [35, Theorem 5.1] for continuous time. A useful consequence in our framework is that for undefaultable strategies, maximality only needs to be tested from time 0, i.e. on \( [0, T] \).

**Corollary 4.11.** Let \( \sigma_1 \leq \sigma_2 \in T_{[0, T]} \) be stopping times. If \( \vartheta \in \mathcal{U}_{\sigma_1} \) is strongly maximal for \( \mathcal{U}_{\sigma_1} \), it is also strongly maximal for \( \mathcal{U}_{\sigma_2} \). Hence any \( \vartheta \in \mathcal{U}_0 \) is strongly maximal for each \( \mathcal{U}_\sigma, \sigma \in T_{[0, T]} \), if and only if it is strongly maximal for \( \mathcal{U}_0 \).

An analogous statement holds for “strong” replaced by “weak”.

**Proof.** If \( \vartheta \in \mathcal{U}_{\sigma_1} \subseteq \mathcal{U}_{\sigma_2} \) fails to be strongly maximal for \( \mathcal{U}_{\sigma_2} \), there is a nonzero contingent claim \( F \) at time \( T \) with \( \Pi_{\sigma_2} V_T(\vartheta) + F \mid \mathcal{U}_{\sigma_2} \) \( \leq V_{\sigma_2}(\vartheta) < \infty \) \( \mathbb{P}\text{-a.s.} \). Proposition 4.10, monotonicity and the definition of superreplication prices then give

\[
\Pi_{\sigma_1} \left( V_T(\vartheta) + F \mid \mathcal{U}_{\sigma_1} \right) = \Pi_{\sigma_1} \left( \Pi_{\sigma_2} \left( V_T(\vartheta) + F \mid \mathcal{U}_{\sigma_2} \right) \bigg| \mathcal{U}_{\sigma_1} \right) \leq \Pi_{\sigma_1} \left( V_{\sigma_2}(\vartheta) \bigg| \mathcal{U}_{\sigma_1} \right) \leq V_{\sigma_1}(\vartheta) \ \mathbb{P}\text{-a.s.}
\]

So \( \vartheta \) fails to be strongly maximal for \( \mathcal{U}_{\sigma_1} \), and we arrive at a contradiction.
5 The main results: Bubbles versus martingales

This section derives the announced connections between interesting bubbly markets and strict local martingales. A large part of the literature on bubbles in financial markets starts, in the classic setup of Example 2.2, with the assumption that the discounted price process $X$ is a strict local martingale (sometimes under $P$ itself, sometimes under a risk-neutral or valuation measure $Q$ chosen in some way). In clear contrast, our definition of an interesting bubbly market yields the conclusion that we must have strict local martingale properties. We formulate this in Theorem 5.5 below and provide more discussion after the result. In a second part, we introduce a notion of no dominance and show that this characterises the difference between dynamic viability and dynamic efficiency.

5.1 Bubbly markets and strict local martingales

Recall that $S$ is a bubbly market by definition if it is not dynamically efficient, i.e. if some invest-and-keep strategy can be improved (approximately) by dynamic trading. To characterise bubbly markets via dual objects and martingale properties, we first give some invest-and-keep strategy can be improved (approximately) by dynamic trading. To characterise bubbly markets via dual objects and martingale properties, we first give

**Lemma 5.1.** Suppose $S$ is dynamically viable and fix a stopping time $\sigma \in T_{[0,T]}$. Then $\vartheta \in U_\sigma(S)$ is weakly maximal for $U_\sigma(S)$ if and only if it is strongly maximal for $U_\sigma(S)$.

**Proof.** Strong clearly implies weak maximality.

Conversely, let $\vartheta \in U_\sigma$ be weakly maximal. We first claim that for each $\bar{\vartheta} \in U_\sigma$ with $V_T(\bar{\vartheta}) \geq V_T(\vartheta)$ $P$-a.s., we have $V(\bar{\vartheta}) \geq V(\vartheta)$ $P$-a.s. on $[\sigma,T]$, so that $\bar{\vartheta} - \vartheta \in U_\sigma$. Indeed, if $\tau \in T_{[\sigma,T]}$ is a stopping time such that $A := \{V_\tau(\bar{\vartheta}) < V_\tau(\vartheta)\}$ has $P[A] > 0$, we take a numéraire strategy $\eta$ and set $\hat{\vartheta} := \vartheta I_{[\tau,T]} + (I_A \vartheta + I_A(\bar{\vartheta} + V_\tau(\vartheta - \bar{\vartheta})(S^{(n)}(\eta)))I_{[\tau,T]}$. Then $\hat{\vartheta} \in U_\sigma$, we have $V_\sigma(\hat{\vartheta}) = V_\sigma(\vartheta)$ $P$-a.s., and using that $V_T(\bar{\vartheta}) \geq V_T(\vartheta)$ $P$-a.s. gives

$$V_T(\hat{\vartheta}) = I_A V_T(\vartheta) + I_A \left( V_T(\bar{\vartheta}) + V_\tau(\vartheta - \bar{\vartheta})(S^{(n)}(\eta))V_T(\eta) \right)$$

$$\geq I_A V_T(\vartheta) + I_A \left( V_T(\bar{\vartheta}) + V_\tau(\vartheta - \bar{\vartheta})(S^{(n)}(\eta))V_T(\eta) \right)$$

$$= V_T(\vartheta) + I_A V_\tau(\vartheta - \bar{\vartheta})(S^{(n)}(\eta))V_T(\eta) \quad P\text{-a.s.}$$

Since $P[A] > 0$ and $V_T(\hat{\vartheta}) > V_T(\vartheta)$ on $A$ by the definition of $A$, this shows that $\vartheta$ fails to be weakly maximal, and we arrive at a contradiction which proves our claim.

To show that $\vartheta$ is strongly maximal, suppose to the contrary that there is a nonzero contingent claim $F$ at time $T$ with $\Pi_\sigma(V_T(\vartheta) + F | U_\sigma) \leq V_\sigma(\vartheta)$ $P$-a.s. Take $\delta > 0$ and a strictly positive contingent claim $C$ at time $\sigma$. Then by Lemma 4.9, there exists $\hat{\vartheta} \in U_\sigma$ with $V_T(\hat{\vartheta}) \geq V_T(\vartheta) + F$ $P$-a.s. and $V_\sigma(\hat{\vartheta}) \leq \Pi_\sigma(V_T(\vartheta) + F | U_\sigma) + \delta C \leq V_\sigma(\vartheta) + \delta C$ $P$-a.s. By the first step, $\vartheta' := \vartheta - \hat{\vartheta}$ is in $U_\sigma$. Moreover, $V_T(\vartheta') \geq F$ $P$-a.s. and $V_\sigma(\vartheta') \leq \delta C$ $P$-a.s. so that we get $\Pi_\sigma(F | U_\sigma) \leq \delta C$ $P$-a.s. Letting $\delta \to 0$ gives $\Pi_\sigma(F | U_\sigma) = 0$ $P$-a.s. and so 0 is not strongly maximal for $U_\sigma$, in contradiction to dynamic viability of $S$.  

25
Our first characterisation of dynamic efficiency now follows by combining several results from Herdegen [23]. Recall that due to (2.1), \( S \) is a numéraire market.

**Theorem 5.2.** The following are equivalent:

1) \( S \) is dynamically efficient.

2) For each bounded numéraire strategy \( \eta \), there exists \( Q \approx P \) on \( F_T \) such that \( S^{(\eta)} \) is a (true) \( Q \)-martingale.

3) There exists a pair \((\eta, Q)\), where \( \eta \) is a numéraire strategy and \( Q \approx P \) on \( F_T \), such that \( S^{(\eta)} \) is a (true) \( Q \)-martingale.

4) There exist a representative \( \bar{S} \in S \) and \( Q \approx P \) on \( F_T \) such that \( \bar{S} \) is a (true) \( Q \)-martingale.

**Proof.** Since \( S \) is nonnegative, both the market portfolio \( \eta^S = 1 \) and the corresponding representative \( S^{(\eta^S)} = S/(\sum_{i=1}^N S^i) \) are bounded, as required for Corollaries 4.15 and 4.16 in [23]. Now if we have 1), then \( \eta^S \) is strongly maximal for \( U_0 \) and 2) follows from Corollaries 4.16 (c) \( \Rightarrow \) (a), and 4.15 (a) \( \Rightarrow \) (c), in [23]. Since there exists a bounded numéraire strategy, it is clear that 2) implies 3), and that 3) implies 4).

Suppose we have 4). Fix any stopping time \( \sigma \in \mathcal{T}_{[0,T]} \) and any \( \vartheta \in \mathcal{hU}_\sigma \). As \( \bar{S} \) is a \( Q \)-martingale on \( [\sigma,T] \), so is \( V(\vartheta)(\bar{S}) \). For any \( \vartheta \in \mathcal{U}_\sigma \), the process \( V(\vartheta)(\bar{S}) \) is a nonnegative stochastic integral of a \( Q \)-martingale and hence a \( Q \)-supermartingale. So if \( V_T(\vartheta) \geq V_T(\bar{\vartheta}) \), we get \( V_\sigma(\vartheta)(\bar{S}) \leq V_\sigma(\bar{\vartheta})(\bar{S}) \), and this shows that \( \vartheta \) is weakly maximal for \( U_\sigma \). But 4) implies by Theorem 4.3 also that \( S \) is dynamically viable, and so weak maximality in \( U_\sigma \) is equivalent to strong maximality in \( U_\sigma \), by Lemma 5.1. Hence we get 1) and the proof is complete.

With the help of the above dual characterisation, we can give some equivalent primal descriptions of dynamically efficient markets.

**Corollary 5.3.** The following are equivalent:

1) \( S \) is dynamically efficient.

2) The market portfolio \( \eta^S = 1 \) (buy and hold one unit of each asset) is strongly maximal for \( U_0(S) \).

3) The strategy \( e^i \) (buy and hold one unit of asset \( i \)) is strongly maximal for \( U_0(S) \), for each \( i = 1, \ldots, N \).

4) For each stopping time \( \sigma \in \mathcal{T}_{[0,T]} \), each \( \vartheta \in \mathcal{bU}_\sigma(S) \) is strongly maximal for \( U_\sigma(S) \).

**Proof.** 4) implies 1) since \( \mathcal{bU} \subseteq \mathcal{bU} \), and 1) trivially implies 2), which is in turn equivalent to 3) by [23, Corollary 4.16]. To see that 3) implies 4), we first argue as in [23, Lemma 4.13] to show that \( V(\vartheta)(S^{(\eta^S)}) \) is a true martingale on \([\sigma,T]\) for \( \vartheta \in \mathcal{bU}_\sigma \), and then proceed as in the second part of the proof of Theorem 5.2, with \( \vartheta \in \mathcal{bU}_\sigma \) there replaced by \( \vartheta \in \mathcal{bU}_\sigma \). 

An immediate consequence from Corollary 5.3 is
Corollary 5.4. \( S \) is a bubbly market if and only if there exists an index \( i \in \{1, \ldots, N\} \) such that the invest-and-keep strategy \( e^i \) for asset \( i \) is not strongly maximal for \( \mathcal{U}_0(S) \).

In words, \( S \) is bubbly if and only if it contains a primary asset which is so bad (or stupid) that it can be beaten, without going into debt, by dynamic trading in the market.

In the classic setup with \( S = (1, X) \), Corollary 5.4 shows that a bubbly market in our sense can arise in two ways. Perhaps one of the risky assets in \( X \) can be dominated by dynamic trading in the other risky assets and the bank account; then that risky asset has a bubble. But alternatively, the bank account itself may be dominated by trading in the other (risky) assets, and in that case, choosing it a priori as numéraire was rash—discounting with such a bank account is not a good idea from an economic perspective.

Corollary 5.4 gives a characterisation of bubbly markets, but not in terms of any local martingale properties. This is very natural—a market can be bubbly (i.e. not dynamically efficient) simply because it admits some arbitrage (i.e. is not dynamically viable), and then we of course cannot expect any martingale properties at all. But if we eliminate that case, the situation changes. Recall again that \( S \) satisfies (2.1).

Theorem 5.5. The following are equivalent:

1) \( S \) is an interesting bubbly market.

2) The zero strategy is strongly maximal in \( \mathcal{U}_0 \), and \( S \) is economically bubbly.

3) There exist a numéraire strategy \( \eta \) and \( Q \approx P \) on \( \mathcal{F}_T \) such that \( S^{(\eta)} \) is a local \( Q \)-martingale; and for any such pair \((\eta, Q)\), the process \( S^{(\eta)} \) is a strict local \( Q \)-martingale.

4) There exist \( \tilde{S} \in S \) and \( Q \approx P \) on \( \mathcal{F}_T \) such that \( \tilde{S} \) is a local \( Q \)-martingale; and for any such pair \((\tilde{S}, Q)\), the process \( \tilde{S} \) is a strict local \( Q \)-martingale.

Proof. If \( S \) is an interesting bubbly market, it is by definition dynamically viable. So the zero strategy is strongly maximal in \( \mathcal{U}_0 \) for each \( \sigma \in \mathcal{T}_{[0,T]} \), or equivalently in \( \mathcal{U}_0 \), by Corollary 4.11. The equivalence of 1) and 2) therefore follows from Proposition 2.19.

By Theorem 4.3, dynamic viability of \( S \) as in 1) is equivalent to the first properties in 3) or 4). In combination with Theorem 5.2, dynamic efficiency of \( S \) then fails as in 1) if and only if the second properties in 3) or 4) hold. This completes the proof.

Theorem 5.5 can be seen as the main result of our paper. It shows that our concept of an interesting bubbly market is equivalently characterised by the appearance of strict local martingales. Moreover, this result is robust in the sense that for each market representative \( \tilde{S} \), we have the strict local martingale property simultaneously under all possible valuation measures \( Q \)—it cannot happen that we “see” a bubble under one measure and no bubble under another. The reason is that our definition of an interesting bubbly market is formulated in terms of superreplication prices. Like these, our bubble concept does not depend on the a priori choice of a valuation or risk-neutral or martingale measure \( Q \). This is in marked contrast to the approach of Protter et al. [30, 31, 46] where the way in which “the market has chosen \( Q \)” does not become completely clear (at least, to us).
Remark 5.6. Note that we are in a market with $N > 1$ traded primary assets. Saying that a representative $\bar{S}$, which is an $\mathbb{R}^N$-valued process, is a strict local $\mathbb{Q}$-martingale means that there is at least one coordinate $\bar{S}^i$ with $i \in \{1, \ldots, N\}$ which has the local, but not the true $\mathbb{Q}$-martingale property. This reflects Corollary 5.3 which says that the market fails to be dynamically efficient if and only if at least one of the invest-and-keep strategies $e^i$ is not strongly maximal for $\mathcal{U}_0$.

5.2 No dominance and true martingales

From Definition 2.14 and the proof of Proposition 2.19, we can see that understanding interesting bubbly markets depends crucially on the difference between dynamic efficiency and dynamic viability. In this section, we characterise this difference with the concept of no dominance, which in itself has some history. It is folklore in mathematical finance that simple risk-neutral valuation results need something extra in addition to absence of arbitrage. This important insight goes back to R. Merton [42] who wrote that “a necessary condition for a rational option pricing theory is that the option be priced such that it is neither a dominant nor a dominated security” and explained that “security (portfolio) $A$ is dominant over security (portfolio) $B$ if on some known date in the future, the return on $A$ will exceed the return on $B$ for some possible states of the world, and will be at least as large as on $B$, in all possible states of the world”.

The above formulation is intuitive, but not very precise. Neither “security” or “portfolio” nor “return” are exactly defined. Subsequent papers have developed different mathematical formulations for the idea, and the key difference lies precisely in those two terms.

The works of Protter and co-authors [30, 31, 46] incorporate “return” by the assumption that each financial product (including basic assets and dynamic trading strategies) has a market price at each time. They do not explain where this comes from; results are obtained by imposing certain structural assumptions on market prices, including “no dominance”. In contrast, Jarrow/Larsson [29] only talk about basic assets and compute the “return” from the value processes of self-financing strategies. This is more specific than the approach in [30, 31, 46], but it also gives in our view potentially sharper results with weaker assumptions on the underlying market. In particular, one can try to impose “no dominance” only on basic assets and then try to deduce analogous properties for suitable valuations applied to complex assets, portfolios or derivatives. We therefore follow [29] in spirit when we introduce our numéraire-independent versions of no dominance.

Definition 5.7. The market $\mathcal{S}$ is said to satisfy

- **static no dominance** if the market portfolio $\eta^\mathcal{S} = 1$ is weakly maximal for $h\mathcal{U}_\sigma(\mathcal{S})$, for each $\sigma \in \mathcal{T}_{[0,T]}$.

- **dynamic no dominance** if the market portfolio $\eta^\mathcal{S} = 1$ is weakly maximal for $\mathcal{U}_\sigma(\mathcal{S})$, for each $\sigma \in \mathcal{T}_{[0,T]}$.

Due to Corollary 4.11, dynamic no dominance is equivalent to the market portfolio $\eta^\mathcal{S}$ being weakly maximal for $\mathcal{U}_0$. Moreover, one can show that the latter holds if and only if for $i = 1, \ldots, N$, the invest-and-keep strategies $e^i$ for each primary asset $i$ are weakly...
maximal for $\mathcal{U}_0$. In fact, the “only if” part is clear since any improvement of an $e^i$ will also improve $\eta^S$, and the “if” part follows from [23, Corollary 3.8], which proves that the weakly maximal strategies in $\mathcal{U}_0$ form a convex cone. This shows that our definition of dynamic no dominance is very close in spirit to the concept of no dominance in Jarrow/Larsson [29]. On the other hand, the concept of static no dominance seems to be new. It is more delicate to analyse; we mention for example that static no dominance is not equivalent to weak maximality of $\eta^S$ (or of $e^i$, for $i = 1, \ldots, N$) for $\mathcal{U}_0$. This can be easily seen if we take Example 3.6 and modify it slightly so that the two possible values of $X_1$ at time 1 are no longer 2 and 1, but 2 and 3/2. We leave the details to the reader.

Our next result connects the notions introduced so far. It shows that no dominance is precisely the extra ingredient that distinguishes efficiency from viability.

**Proposition 5.8.** $S$ is dynamically efficient if and only if it is dynamically viable and satisfies dynamic no dominance.

**Proof.** Efficiency trivially implies viability and yields, in the dynamic case, that $\eta^S$ is strongly (and a fortiori weakly) maximal for $\mathcal{U}_\sigma$, for each $\sigma \in \mathcal{T}_{[0,T]}$. Conversely, we know from Lemma 5.1 that under dynamic viability, weak is equivalent to strong maximality of $\eta^S$ for $\mathcal{U}_0$. So dynamic efficiency follows from Corollary 5.3.

**Remark 5.9.** One can also prove the static analogue of Proposition 5.8 where “dynamic” is replaced by “static” in all three appearances. The arguments are a bit different (see [22, Proposition VIII.3.19]) and we omit them for reasons of space. However, we mention that this argument actually shows that static efficiency and static no dominance are equivalent.

One important result in the classic setup is that no dominance is the extra strengthening of “absence of arbitrage” required to obtain the existence of an equivalent true (as opposed to local or $\sigma$-) martingale measure; see Theorem 3.2 of [29]. Our next result establishes the same connection in our numéraire-independent framework.

**Corollary 5.10.** The following are equivalent:

1) $S$ satisfies NINA and dynamic no dominance.

2) There exists a pair $(\eta, Q)$, where $\eta$ is a bounded numéraire strategy and $Q \approx P$ on $\mathcal{F}_T$ is such that the $V(\eta)$-discounted price process $S^{(\eta)} = \frac{S}{V(\eta)}$ is a $Q$-martingale.

**Proof.** By Theorem 4.3, NINA or strong maximality of 0 in $\mathcal{U}_0$ is equivalent to dynamic viability of $S$. Together with dynamic no dominance, this is by Proposition 5.8 equivalent to dynamic efficiency of $S$, and this in turn is equivalent to 2) by Theorem 5.2.

**6 Further examples**

Due to Proposition 4.2, there are no interesting bubbly markets in finite discrete time. This changes in continuous time.
Example 6.1 (Complete interesting bubbly markets). Let $S$ be the market generated by $S = (S^1, S^2) = (X, 1)$, where $X$ is a strict local $\mathbb{P}$-martingale; so the bank account is asset 2, for a change. We suppose that $S$ is complete, which means that $X$ has the predictable representation property in the filtration $(\mathcal{F}_t)_{t \in [0,T]}$ we work with. We claim that $S$ is dynamically viable, but not dynamically efficient; so this is a generic example of an interesting bubbly market.

First, dynamic viability follows directly from Theorem 4.3 when we choose for $\eta$ the invest-and-keep strategy $e^2 = (0, 1)$ of the second asset. Next, completeness and continuity of $X$ imply that the only $\mathbb{P}$-martingales strongly $\mathbb{P}$-orthogonal to $X$ are constants, since $\mathcal{F}_0$ is $\mathbb{P}$-trivial. Thus the density process $Z$ of any ELMM $Q$ for $X$ must be constant, hence 1, so that $Q \equiv \mathbb{P}$. So because $X$ is a strict local $\mathbb{P}$-martingale, there cannot be any $Q \approx \mathbb{P}$ which makes $X$ a true $Q$-martingale, and since this means that (c) in Theorem 5.2 with $\eta = e^2$ fails, we conclude that $S$ is not dynamically efficient.

For applications to valuation, it is also of interest to study when $S$ is even a nontrivial bubbly market. For that, we only need to check if it is addition statically efficient. One can show (see [22, Theorem VIII.3.16]) that this is here the case if and only if

$$X_s \in \text{ri conv supp } \mathcal{L}(X_T | \mathcal{F}_s) \ \text{ $\mathbb{P}$-a.s. for each } s \in [0,T),$$

(6.1)

where $\text{supp } \mathcal{L}(S^1_T | \mathcal{F}_s)$ is the ($\omega$-dependent) support of the regular conditional distribution of $X_T$ given $\mathcal{F}_s$, $\text{conv}$ denotes the convex hull, and $\text{ri}$ the relative interior.

For a complete market as above, there is no essential difference between our notion of bubbliness and the definition via strict local martingales. This is not surprising because most subtleties appear only when there is no unique candidate for a valuation measure.

Example 6.2. For a concrete example of a strict local $\mathbb{P}$-martingale which has the predictable representation property, we can go back to Example 3.3 where $S = (1, X)$ and $X$ is under $\mathbb{P}$ a $\text{BES}^\beta$ process. The representative $S^{(e^2)} = (1/X, 1) =: (Y, 1)$ is then of the form that we want. More generally, we could assume that $Y$ is a constant elasticity of variance (CEV) process, i.e., satisfies the SDE, with $\sigma > 0$ and $\beta > 1$,

$$dY_t = \sigma |Y_t|^\beta \ dW^\mathbb{P}_t, \quad Y_0 = y_0 > 0.$$ (6.2)

It is well known\footnote{It seems difficult to find an exact single reference which rigorously shows that the SDE (6.2) has a unique strong solution and that this is a strict local $\mathbb{P}$-martingale. But of course this follows easily from the general theory of one-dimensional SDEs and the explicit transition density computed by Emanuel and MacBeth [13, Equation (7)].} that the SDE (6.2) has a unique strong solution $Y$ which is a positive continuous strict local $\mathbb{P}$-martingale; see [37, Section 9.8] for a detailed discussion of the CEV model. As in Example 3.3 and Remark 3.4, one can also see that the CEV process has the predictable representation property for its own filtration (which can equivalently be generated by $W^\mathbb{P}$). The $\text{BES}^2$ process in (3.1) is the special case where $\beta = 2, \sigma = 1$.

For this example, one can check (6.1) by using the transition densities $f(T, y; s, x)$ for the conditional distribution at time $T$, given that we are in $x$ at time $s$. The explicit formula for the CEV model can be found e.g. in Emanuel/MacBeth [13, Equation (7)].
Example 6.3 (An incomplete interesting bubbly market). On \([0,T]\), consider two independent \(P\)-Brownian motions \(W^P\) and \(W'\) with respect to a given filtration \((\mathcal{F}_t)_{t \in [0,T]}\), which need not be generated by \((W^P, W')\). The market \(S\) is generated by \(S = (S^1, S^2) = (X, 1)\), where \(X\) satisfies the SDE

\[
dX_t = V_t |X_t|^\beta dW^P_t, \quad X_0 = x_0 > 0. \tag{6.3}
\]

Here \(\beta > 1\) is a constant and the stochastic volatility \(V = (V_t)_{t \in [0,T]}\) satisfies the SDE

\[
dV_t = \alpha(V_t - \sigma)(V_t - \sigma) dW'_t, \quad V_0 = v_0 \in (\sigma, \sigma), \tag{6.4}
\]

for constants \(\alpha > 0\) and \(\sigma > \sigma > 0\). This can be interpreted as a CEV model (see Example 6.2) with stochastic volatility \(V\) and elasticity of variance \(\beta > 1\). It is not difficult to check that (6.4) has a unique strong solution satisfying \(\sigma < V < \sigma\) \(P\)-a.s.; see e.g. Rady [47, Section 3]. We remark that the exact form of \(V\) is not important for the argument that follows; we only use that \(V\) is a continuous \((\mathcal{F}_t)\)-adapted strong Markov process uniformly bounded from above and below by positive constants.

To argue that (6.3) has a unique strong solution, we first show a more general result: If \(Q \approx P\) on \(\mathcal{F}_T\) is such that \(W^P\) is a \(Q\)-Brownian motion on \([0,T]\), then (6.3) has a unique strong solution \(X\) with \(E_Q[X_t] < x_0\), \(t \in (0,T]\), i.e., \(X\) is a strict local \(Q\)-martingale. Moreover, there exists \(\varepsilon \in (0,T]\) which depends on \(x_0\), but not on \(v_0\), such that

\[
E_Q[X_{\varepsilon}] > \frac{x_0}{2}. \tag{6.5}
\]

Let us argue these claims, using [45, Chapter V] as reference. A solution to (6.3) under \(Q\) (up to a possible explosion time) is unique because \(f : [0,\infty) \times \Omega \times [0,\infty) \to \mathbb{R}, f(t,\omega,x) := V_t(\omega)|x|^{\beta}\), is uniformly in \(t\) locally random Lipschitz in \(x\), i.e., for each \(n \in \mathbb{N}\), there is a finite random variable \(K_n\) with \(\sup_{t \in [0,T]}|f(t,\omega,x) - f(t,\omega,y)| \leq K_n(\omega)|x-y|\) for all \(x,y \in [0,n]\). To establish existence and prove the remaining assertions, we use a time-change argument reducing (6.3) to the SDE of the standard CEV model. To simplify the presentation, we assume that after possibly enlarging the original probability space, there exists a \(Q\)-Brownian motion \((W^Q_t)_{t \geq 0}\) with \((W^Q_t)_{t \in [0,T]} = W^P\). Denote by \(\mathcal{N}\) the \(Q\)-nullsets in \(\mathcal{F}_T\) \(\vee \sigma(W^Q_s; s \geq 0)\) and set \(\tilde{\mathcal{F}}_t = \mathcal{F}_{t\wedge T} \vee \sigma(W^Q_s; s \leq t) \vee \mathcal{N}\) for \(t \geq 0\). Then \((\tilde{V}_t)_{t \geq 0}\) defined by \(\tilde{V}_t = V_{t\wedge T}\) is a continuous \((\tilde{\mathcal{F}}_t)\)-adapted process which satisfies \((\tilde{V}_t)_{t \in [0,T]} = V\) and takes values in \((\sigma, \sigma)\) \(Q\)-a.s. We are going to construct a strong solution on \([0,\infty)\) of the SDE

\[
d\tilde{X}_t = \bar{V}_t |\tilde{X}_t|^\beta dW^Q_t, \quad \tilde{X}_0 = x_0 > 0, \tag{6.6}
\]

and it is clear that \(X_t = \tilde{X}_t\) for \(t \in [0,T]\) is then a strong solution to (6.3).

Define \(\tilde{M}_t = \int_0^t \tilde{V}_s dW^Q_s\) and \(\Lambda_t = \int_0^t |\tilde{V}_s|^2 ds\) for \(t \geq 0\). Then \(\tilde{M}\) is under \(Q\) a continuous local \((\tilde{\mathcal{F}}_t)\)-martingale null at 0, and \(\Lambda\) has \(Q\)-a.s. continuous trajectories, is null at 0, strictly increasing, and satisfies \(Q\)-a.s.

\[
\bar{\sigma}^2 t > \Lambda_t > \sigma^2 t \quad \text{for} \ t \geq 0. \tag{6.7}
\]

Define \(\tau = \inf\{s \geq 0 : \Lambda_s \geq t\}\) for \(t \geq 0\) so that \(\tau\) is an increasing continuous time change for \((\tilde{\mathcal{F}}_t)_{t \geq 0}\). Define \(\tilde{\mathcal{F}}_t := \tilde{\mathcal{F}}_{\tau_t}\) and \(\tilde{W}_t := \tilde{M}_{\tau_t}\) for \(t \geq 0\). Then \(\tilde{W}\) is under \(Q\)
a continuous local \((\tilde{\mathcal{F}}_t)\)-martingale with \(\langle \tilde{W} \rangle_t = \langle \tilde{M} \rangle_t = \Lambda_t = t\) Q-a.s. and hence a Q-Brownian motion for \((\tilde{\mathcal{F}}_t)_{t \geq 0}\). In this time-changed filtration, consider the SDE for the standard CEV model,

\[
d\tilde{X}_t = |\tilde{X}_t|^\beta d\tilde{W}_t, \quad \tilde{X}_0 = x_0 > 0.
\] (6.8)

This has a unique strong solution \(\tilde{X}\) which is a positive continuous strict local Q-martingale (cf. Example 6.2). Moreover, the explicit formula for the transition density ([13, Equation (7)]) yields \(\lim_{t \to 0} E_Q[\tilde{X}_t] = x_0\). Define \(\tilde{X}_t = \tilde{X}_{\Lambda_t}\) for \(t \in [0, T]\), and note that \(\tilde{M}_t = \tilde{W}_{\Lambda_t}, t \geq 0\). Then \(\tilde{X}\) is a positive continuous local Q-martingale for the filtration \((\tilde{\mathcal{F}}_t)_{t \in [0,T]}\), and plugging in the definitions and using (6.8) shows that it satisfies the SDE

\[
d\tilde{X}_t = |\tilde{X}_t|^\beta d\tilde{M}_t = \tilde{V}_t|\tilde{X}_t|^\beta dW_t^Q, \quad \tilde{X}_0 = x_0,
\]

as desired for (6.6). Moreover, \(\tilde{X}\) is under \(Q\) a positive \((\tilde{\mathcal{F}}_t)\)-supermartingale by Fatou’s lemma, and so by (6.7) and the properties of \(\tilde{X}\),

\[E_Q[\tilde{X}_t] = E_Q[\tilde{X}_{\Lambda_t}] \leq E_Q[\tilde{X}_{\tilde{a}^2_t}] < x_0.\]

By the same argument, \(E_Q[\tilde{X}_t] \geq E_Q[\tilde{X}_{\tilde{a}^2_t}]\), and since the right-hand side does not depend on \(\tilde{a}_0\), this together with \(\lim_{t \to 0} E_Q[\tilde{X}_t] = x_0\) establishes (6.5).

To show that \(S\) is dynamically viable but fails to be dynamically efficient, we first note that \(S = S(e^1)\) and \(X = V(e^1)(S(e^2))\), where \(e^1 = (1,0)\) and \(e^2 = (0,1)\) are the invest-and-keep strategies of the first and second asset. By the above result for \(Q = P\), \(X\) is a local \(P\)-martingale and so \(S\) is dynamically viable by Theorem 4.3. To show that \(S\) is not dynamically efficient, it suffices by Theorem 5.2 to show that there is no \(Q \approx P\) on \(\mathcal{F}_T\) such that \(X\) is a (true) Q-martingale. But if \(X\) is a local Q-martingale under \(Q \approx P\) on \(\mathcal{F}_T\), then \(W^P = f V^{-1}|X|^{-\beta} dX\), by (6.3) and strict positivity of \(X\) and \(V\), is a continuous local Q-martingale with quadratic variation \(\langle W^P \rangle_t = \int_0^t V_s^{-2}|X_s|^{-2\beta} d\langle X \rangle_s = t, t \in [0,T]\), and so \(W^P\) is also a Q-Brownian motion. Again by the above result, \(X\) is therefore a strict local Q-martingale, and so \(S\) is not dynamically efficient.

We remark (without proof; see [22, Example VIII.4.3]) that \(S\) is also statically efficient. Thus \(S\) is not only an interesting, but even a nontrivial bubbly market.

Example 6.3 is of interest for several reasons. First of all, it is a CEV model with stochastic volatility and thus quite realistic from a practical perspective. In fact, if we replace the volatility process \(V\) from the SDE (6.4) by a geometric Brownian motion, we get the well-known SABR model (see [20]). Next, as \(X = S^1\) is a strict local \(P\)-martingale, we have a bubble model in the sense of Loewenstein/Willard [39], Protter et al. [30, 31, 46] or Cox/Hobson [3], among others; see Section 7. However, we actually have more. In most approaches to bubble modelling in incomplete markets (e.g. [31, 46]), one fixes an ELMM or risk-neutral measure \(Q\) for \(S\) and assumes that \(S\) is under \(Q\) a strict local martingale. This should really be called a Q-bubble, because there may well be another ELMM \(Q'\) under which \(S\) is a true martingale; so the above notion of a bubble can depend in a crucial way on the choice of the risk-neutral measure. (For a more thorough discussion of that issue, see Section 7.) In Example 6.3, this cannot happen. Because we have an interesting bubbly market, Theorem 5.5 tells us that for all possible representatives
$S \in \mathcal{S}$ and all ELMMs $Q$, we always have for $S$ under $Q$ a strict local martingale. In other words, Example 6.3 gives a concrete incomplete market with a bubble which is robust towards the choice of the ELMM one wants to use. This can also be seen from the above arguments—we show that $X$ is a strict local $Q$-martingale whenever we have (6.3), (6.4) under some $Q \approx P$ on $\mathcal{F}_T$. Apart from Jarrow/Larsson [29, Theorem 5.7], such a robust bubble model has not been presented in the literature so far. (Note also that in [29, Theorem 5.7], it is not argued that we have strict local $Q$-martingales on a finite time horizon.) Last but not least, the market in Example 6.3 is statically efficient and so invest-and-keep trading is still optimal in its own class. Even though it is neither proved nor discussed here, we point out that this property is crucial when studying economically consistent valuation of contingent claims.

The next example gives a concrete model where $S$ is under some ELMM $Q$ a strict local martingale, but under another ELMM $Q'$ a true martingale. Of course, by Theorem 5.5, the market generated by this model is then not an interesting bubbly market in our sense.

Example 6.4 (A $Q$-bubble which is not a $Q'$-bubble). Start with two $Q$-Brownian motions $W^i = (W^i_t)_{t \in [0,T]}, i = 1, 2$, with respect to a given filtration $(\mathcal{F}_t)_{t \in [0,T]}$; this need not be generated by $(W^1, W^2)$. We assume that $W^1$ and $W^2$ are positively but not perfectly correlated: there is a constant $\lambda \in (0,1)$ such that $d\langle W^1, W^2 \rangle_t = \lambda \, dt$ (we use $\rho$ for something else below). The market $\mathcal{S}$ is generated by $S = (S^1, S^2) = (1, X_t)_{t \in [0,T]}$, where $X$ satisfies the SDE, for some constant $\xi > 0$,

$$dX_t = \xi X_t V_t dW^1_t, \quad X_0 = x_0 > 0.$$  \hfill (6.9)

The volatility process $V = (V_t)_{t \in [0,T]}$ is stochastic and satisfies the SDE, for some $b > 0$,

$$dV_t = b V_t dW^2_t, \quad V_0 = 1.$$  \hfill (6.10)

It is clear that (6.10) and (6.9) have unique strong solutions $V$ and $X$. We claim that

(i) $X$ is a strict local $Q$-martingale on $[0,T]$;

(ii) there is a probability measure $Q' \approx Q$ on $\mathcal{F}_T$ such that $X$ is a true $Q'$-martingale.

To prove this, we use the results of Sin [52]. Setting $a := (b \lambda, b \sqrt{1 - \lambda^2})$, $\sigma := (\xi,0)$ and $\rho = 0$, we are exactly in the setup of [52, Theorems 3.3 and 3.9] with $\alpha = 1$. Note that $a \cdot \sigma = \xi b \lambda > 0$, and $a, \sigma$ are not parallel. So we immediately get the existence of $Q'$ (called $Q^a$ in [52, Theorem 3.9]) for (ii). The result (i) does not follow directly from [52, Theorem 3.3], since a strict local martingale on $[0,\infty)$ might still be a true martingale on a given finite interval. But $X$ is a positive local $Q$-martingale, hence a $Q$-supermartingale, and so it suffices to show that $\mathbb{E}[X_T] < x_0$. For that, by [52, Lemma 4.2], it is enough to show that $Q[\hat{\tau} < T] > 0$, where $\hat{\tau}$ is the explosion time of the SDE

$$d\hat{V}_t = b\hat{V}_t d\hat{W}_t + b\xi \lambda \hat{V}_t^2 \, dt, \quad \hat{V}_0 = 1,$$  \hfill (6.11)

with a generic $Q$-Brownian motion $\hat{W} = (\hat{W}_t)_{t \geq 0}$. For the rest of the example, denote by $\hat{V}$ the canonical process on the path space $C([0,\infty); (0,\infty) \cup \{\Delta\})$, where $\Delta$ is an absorbing cemetery state, by $P_v$ the distribution on the path space of the solution of (6.11) with
initial value $v > 0$, and by $\vartheta$ the shift operator. It follows from [52, Lemma 4.3] that under each $P_v$, $\hat{V}$ explodes in finite time with positive probability and is valued in $(0, \infty)$ before the explosion (the argument in [52] does not depend on the initial value $v$). With $T_v := \inf\{T \geq 0 : P_v[\hat{\tau} < T] > 0\}$ for $v > 0$, this means that $T_v < \infty$. We claim that in fact $T_v = 0$ for all $v > 0$, and this will complete the proof, because we then have $P_v[\hat{\tau} < T] > 0$ for all $T > 0$, as desired.

We first show that $v \mapsto T_v$ is decreasing. Indeed, if $T_{v_1} < T_{v_2}$ for $0 < v_1 < v_2$, there is $\varepsilon > 0$ with $P_{v_1}[\hat{\tau} < T_{v_1} + \varepsilon] > 0$ and $P_{v_2}[\hat{\tau} < T_{v_1} + \varepsilon < T_{v_2}] = 0$. With $\tau_{v_2}^\varepsilon := \inf\{t > 0 : \hat{V}_t > y\}$, we can use $\hat{\tau} = \hat{\tau} \circ \vartheta_{\tau_{v_2}^\varepsilon}$ and the strong Markov property to get

$$0 < P_{v_1}[\hat{\tau} < T_{v_1} + \varepsilon] = P_{v_1}[\hat{\tau} \circ \vartheta_{\tau_{v_2}^\varepsilon} < T_{v_1} + \varepsilon] = E_{v_1}[P_{v_2}[\hat{\tau} < T_{v_1} + \varepsilon]] = 0,$$

a contradiction. So $v \mapsto T_v$ is decreasing and $T_\infty := \lim_{v \to \infty} T_v$ exists in $[0, \infty)$. If $T_\infty > 0$, there is $\varepsilon > 0$ with $P_v[\hat{\tau} \leq \varepsilon] = 0$ for all $v \in (0, \infty)$, and the Markov property gives

$$P_v[\hat{\tau} \leq 2\varepsilon] = P_v[\hat{\tau} \leq 2\varepsilon, \hat{\tau} > \varepsilon] = P_v[\hat{\tau} \circ \vartheta_{\varepsilon} \leq 2\varepsilon] = E_v[P_{V_\varepsilon}[\hat{\tau} \leq \varepsilon]] = 0$$

for all $v \in (0, \infty)$. Iterating this argument yields $P_v[\hat{\tau} \leq n\varepsilon] = 0$ for all $n \in \mathbb{N}$, $v \in (0, \infty)$, and we obtain a contradiction. So $T_\infty = 0$. Finally, we show that $T_v = 0$ for all $v > 0$. If this fails, there is $v_0 \in (0, \infty)$ with $T_{v_0} > 0$, and then there is $\varepsilon > 0$ with $P_{v_0}[\hat{\tau} \leq 2\varepsilon] = 0$. Using that $v \mapsto T_v$ is decreasing and $T_\infty = 0$, pick $v_1 > v_0$ large enough that $T_{v_1} < \varepsilon$; then $P_v[\hat{\tau} \leq \varepsilon] > 0$ for all $v \geq v_1$ since $T_v$ is decreasing in $v$. Because $b\xi \lambda > 0$, a standard comparison argument for SDEs yields $\hat{V} \geq \hat{V}_0 P_{v_0}$-a.s., where $\hat{V} = v_0 \mathcal{E}(b\hat{W})$ satisfies $d\hat{V}_t = b\hat{V}_t d\hat{W}_t$, and so $P_{v_0}[\hat{V}_\varepsilon \geq v_1] \geq P_{v_0}[\hat{V}_\varepsilon \geq v_1] > 0$ since $\hat{V}_\varepsilon$ has a lognormal distribution. Using the Markov property then gives the contradiction

$$0 = P_{v_0}[\hat{\tau} \leq 2\varepsilon] \geq P_{v_0}[\hat{\tau} \leq 2\varepsilon, \hat{V}_\varepsilon \geq v_1, \hat{\tau} > \varepsilon] = P_{v_0}[\hat{\tau} \circ \vartheta_{\varepsilon} \leq \varepsilon, \hat{V}_\varepsilon \geq v_1]$$

$$= E_{v_0}[E_{V_\varepsilon}[\hat{\tau} \leq \varepsilon] I_{\{V_\varepsilon \geq v_1\}}] > 0.$$

So $T_v = 0$ for all $v > 0$, and $X$ is a strict local $Q$-martingale on $[0, T]$, for each $T > 0$.

## 7 Comparison to the literature

We have already mentioned in the introduction that the literature on bubbles is too large and diverse for an overview or detailed discussion. As a more modest aim, we compare our approach to some seminal recent papers from the mathematical finance literature. To that end, it is helpful to provide a unified framework within which different approaches can be analysed. For a critical approach to mathematical models for bubbles, we also refer to the recent paper by Guasoni/Rásonyi [19].

### 7.1 Fundamental values

We start with a time horizon $T > 0$ and a filtered probability space $(\Omega, F, F, P)$ with $F = (F_t)_{t \in [0,T]}$. We describe a dividend-paying asset $(\Delta, Y)$ by its cumulative dividend process $\Delta = (\Delta_t)_{t \in [0,T]}$ and its ex-dividend price process $Y = (Y_t)_{t \in [0,T]}$, both in the
same currency units. (We use here \( Y \) and not \( S \) to make the notation consistent with Example 2.2.) We also include a bank account \( B = (B_t)_{t \in [0,T]} \); so if we hold one unit of the asset over a time interval \((t, u]\), we obtain at time \( u \) a total cashflow or gain of

\[
Y_u - \frac{B_u}{B_t} Y_t + B_u \int_t^u \frac{1}{B_s} d\Delta_s.
\]

Its equivalent discounted back to time \( t \) is

\[
B_t \left( \frac{Y_u}{B_u} - \frac{Y_t}{B_t} + \int_t^u \frac{1}{B_s} d\Delta_s \right) =: B_t \left( X_u - X_t + \int_t^u \frac{1}{B_s} d\Delta_s \right) =: B_t (W_u - W_t),
\]

and the discounted gains process from holding one unit of the asset is therefore

\[
W_t = \frac{Y_t}{B_t} + \int_0^t \frac{1}{B_s} d\Delta_s = X_t + \int_0^t \frac{1}{B_s} d\Delta_s, \quad t \in [0, T].
\] (7.1)

If \( Y \) and \( \Delta \) are multidimensional, we add a superscript \( i \) for \( Y, \Delta \) and \( W \). Without dividends, (7.1) reduces to \( W = Y/B = X \), the \( B \)-discounted asset price.

As in the introduction, we denote by \(*Y_t\) the (undiscounted) fundamental value of the asset \((\Delta, Y)\) at time \( t \). If \(*Y_t \neq Y_t\), it is natural to say that the asset has a bubble. But how do we find or define \(*Y_t\)?

Axiomatically, one reasonable approach is to impose that we have a fundamental value operator which assigns fundamental values to assets or general financial products. Such an operator is usually monotone, and if we add linearity and some mild continuity conditions, it is reasonable to assume that it has the form, for \( 0 \leq t \leq u \leq T \),

\[
\Phi^Z_t(\Delta, Y) = B_t \mathbb{E} \left[ \frac{Z_u Y_u}{Z_t B_u} + \int_t^u \frac{Z_s}{Z_t B_s} d\Delta_s \bigg| \mathcal{F}_t \right]
\]

\[
= \frac{1}{\rho_t} \mathbb{E} \left[ \rho_t Y_u + \int_t^u \rho_s d\Delta_s \bigg| \mathcal{F}_t \right] =: \Phi^i_t(\Delta, Y)
\]

\[
= \frac{B_t}{Z_t} \mathbb{E} \left[ \frac{Z_u X_u}{Z_t} + \int_t^u \frac{Z_s}{B_s} d\Delta_s \bigg| \mathcal{F}_t \right],
\] (7.2)

where the positive adapted processes \( Z \) and \( \rho := Z/B \) are often called a deflator or a state price density, respectively. For the “bank account” asset \((0, B)\), we get

\[
\Phi^Z_0(0, B) = B_t \mathbb{E} \left[ \frac{Z_u}{Z_t} \bigg| \mathcal{F}_t \right].
\]

If the bank account has no bubble, \( Z \) is a positive \( \mathbb{P} \)-martingale; more generally, \( Z \) is assumed to be a positive local \( \mathbb{P} \)-martingale with \( Z_0 = 1 \). Similarly, if (7.2) equals \( Y_t \), so that the asset \((\Delta, Y)\) has no bubble, the process \( \rho Y + \int \rho d\Delta = ZX + \frac{Z}{B} d\Delta \) is a \( \mathbb{P} \)-martingale; more generally, this process is assumed to be a local \( \mathbb{P} \)-martingale.

In the case where \( Z \) is a martingale, we can define a probability measure \( Q \) equivalent to \( \mathbb{P} \) by \( dQ = Z_T d\mathbb{P} \) and then rewrite (7.2) as

\[
\Phi^Z_t(\Delta, Y) = B_t \mathbb{E}_Q \left[ \frac{Y_u}{B_u} + \int_t^u \frac{1}{B_s} d\Delta_s \bigg| \mathcal{F}_t \right] = B_t \mathbb{E}_Q \left[ X_u + \int_t^u \frac{1}{B_s} d\Delta_s \bigg| \mathcal{F}_t \right] =: \Phi^Q_t(\Delta, Y)
\]

35
by the Bayes rule. Using (7.1), we can also reformulate this as

$$
\Phi_t^Q(\Delta, Y) = B_t E_Q[W_u - W_t | \mathcal{F}_t] + Y_t.
$$

(7.3)

So if we decide that the fundamental value of $(\Delta, Y)$ at time $t$ is $^* Y_t := \Phi_t^Q(\Delta, Y)$, having a bubble means that $W$ (or $X = Y/B$ in the absence of dividends) is not a $Q$-martingale. In the same way, if $Z$ is not assumed to be a $P$-martingale and we set $^* Y_t := \Phi_t^Z(\Delta, Y)$, having a bubble means that $\rho Y + \int \rho d\Delta = ZX + \int \frac{Z}{B} d\Delta$ is not a $P$-martingale. Note that this depends on $Z$ or $Q$; so one really should talk about a $Q$-bubble or a $Z$-bubble.

Instead of the axiomatic and linear fundamental value operator, one can use the superreplication price for the asset $(\Delta, Y)$, which is defined by

$$
\Psi_t(\Delta, Y) := \text{ess inf}\{v_t \in L^0_+ : \exists \text{ self-financing strategy in } (\Delta, Y)
\text{ with initial wealth } v_t \leq Y_t \text{ at time } t
\text{ and final wealth } V_T \geq Y_T + B_T \int_b^t \frac{1}{B_s} d\Delta_s\}.
$$

By the fundamental hedging duality going back to Kramkov [36] and extended in [16, 53, 23], this can be written as

$$
\Psi_t(\Delta, Y) = \text{ess sup}\{\Phi_t^Z(\Delta, Y) = \Phi_t^\rho(\Delta, Y) : Z > 0 \text{ is a local } P\text{-martingale with } Z_0 = 1
\text{ and such that } \rho Y + \int \rho d\Delta = ZX + \int \frac{Z}{B} d\Delta
\text{ is a local } P\text{-martingale}\}.
$$

(7.4)

In the case of a complete (and suitably arbitrage-free) model, $Z$ or $\rho$ are unique, and in that case, $\Psi_t(\Delta, Y) = \Phi_t^Z(\Delta, Y) = \Phi_t^\rho(\Delta, Y)$ so that both approaches coincide. For incomplete markets, however, this is no longer true; see Example 3.7.

We remark that a process $Z$ as above has been called a local martingale density or a local martingale deflator in the recent literature, and that its existence, for a setting $\Delta \equiv 0$ without dividends, has been shown to be equivalent to the absence-of-arbitrage condition of no unbounded profit with bounded risk (NUPBR); see e.g. Kardaras [33] or Takaoka/Schweizer [54]. This condition, as shown by Kardaras [33], is in turn equivalent to the absence of arbitrage of the first kind (NA1) or absence of cheap thrills, and this latter condition appears also in Loewenstein/Willard [38]; see Herdegen [23] for a discussion of those concepts from a numéraire-independent perspective. For the discussion of bubble modelling, these remarks are at present tangential, but we come back to them later.

With this terminology, we now discuss some important papers from the literature.

### 7.2 Bubbles and equilibrium

In two seminal papers, Loewenstein and Willard [39, 40] start with a financial market $(B,Y,\Delta)$ where $B$ and $Y$ are positive Itô processes in a Brownian filtration, $\Delta \geq 0$ is increasing and $Y, \Delta$ are multidimensional. They assume that there exists a local $P$-martingale $Z > 0$ with $Z_0 = 1$ such that deflated gains $\frac{Z}{B} Y + \int \frac{Z}{B} d\Delta =: \rho Y + \int \rho d\Delta$ form a local $P$-martingale. In [39] but not in [40], they also impose completeness of the market by assuming that $Z$ or $\rho$ is unique. The main goal and result in both papers is a study of the additional restrictions on bubbles that result from market clearing in equilibrium.
Interestingly, there is a shift from [39] to [40] in the definition of fundamental values. In the first paper [39], Loewenstein and Willard use the (linear) definition (7.2) via \( \Phi^0_t(\Delta, Y) \); in the second paper [40], they use the (nonlinear) superreplication price \( \Psi_0(\Delta, Y) \) at time 0 from (7.4) and say that this is “neoclassical economics”. (This actually already appears earlier, in unpublished draft versions of [40] which contain additional results and discussions.) Both papers emphasise that their definition is in line with the previous economic literature (e.g. Diba/Grossman [10], Tirole [55], or Santos/Woodford [49]).

To compare this with our approach, suppose for simplicity that there are no dividends so that \( \Delta \equiv 0 \), and write again \( S \) for a representative of a market \( \mathcal{S} \) in our sense. Let \( \mathcal{S} \) be generated by \( S = (B, Y) \), where \( B \) is a bank account and \( Y \) a stock in the sense of [39, 40]. In our setup, completeness translates into saying that for any numéraire strategy \( \eta \), there is at most one equivalent local martingale measure (ELMM) \( Q \) for \( V(\eta) \)-discounted prices \( S^{(\eta)} \). If \( \mathcal{S} \) is dynamically viable, there exists a pair \((\eta, Q)\) such that \( Q \) is an ELMM for \( S^\eta = S/V(\eta)(S) \) (Theorem 4.3). If \( \mathcal{S} \) is in addition an interesting bubbly market, the (unique, under completeness) ELMM \( Q \) for \( S^{(\eta)} \) is such that \( S^{(\eta), j} \) is a strict local \( Q \)-martingale for at least one index \( j \) (Theorem 5.2 and Corollary 5.4). If \( Z^Q \) is the density process of \( Q \) with respect to \( P \), then \( Z^Q S^{(\eta)} \) is a local \( P \)-martingale and \( Z^Q S^{(\eta), j} \) is a strict local \( P \)-martingale. Setting \( Z := Z^Q B/V(\eta)(S) \), we get that \( Z \) and \( \frac{Z}{B} Y \) are local \( P \)-martingales, and either \( Z \) (if \( j = 1 \)) or \( \frac{Z}{B} Y^i \) (for \( i = j - 1 \) if \( j > 1 \)) is a strict local \( P \)-martingale.

So under completeness, if \( \mathcal{S} \) generated by \( S = (B, Y) \) is an interesting bubbly market, either one of the stocks \( Y^i \) or the bank account \( B \) has a bubble in the sense of [39]. It is not difficult to check that also the converse is true.

In the incomplete case, [40] is also close to our approach, with one difference: The authors only look at the fundamental value at time 0 and use this to decide whether or not the asset has a bubble. We explain below in Section 7.4 why this distinction matters.

### 7.3 Bubbles and mathematics

Jarrow, Protter and Shimbo provide a detailed study of asset price bubbles in two papers—one for complete [30] and one for incomplete markets [31]. Their setup is a financial market \((1, X, \Delta)\) as above, with \( B \equiv 1 \) and \( X, \Delta \geq 0 \) one-dimensional semimartingales. They work on a right-open stochastic interval \([0, \tau]\) with a stopping time \( \tau \) and add a liquidation value \( X_\tau \) at \( \tau \) to the (discounted) stock \( X \), but this is a minor detail; we can replace \( X_\tau \) by the final stock price \( X_T \) without changing the essence of the model. (We point out that the dividend process \( \Delta \) should be increasing for some of the arguments to work.) Instead of the existence of (a local martingale density) \( Z \) (or equivalently NUPBR, as discussed above in Section 7.1), [30, 31] impose the stronger condition NFLVR for the gains process \( W = X + \Delta \); so there exists an ELMM \( Q \) for \( W \) by the fundamental theorem of asset pricing. The first paper [30] on complete markets assumes that \( Q = Q^* \) is unique; the second [31] does not, and we denote by \( \mathcal{M}_c(W) \) the nonempty set of ELMMs \( Q \) for \( W \).

For the complete case [30], the fundamental value is defined as in (7.3) by

\[
X^*_t := \Phi^Q_t(\Delta, X) := \mathbb{E}^Q[X_T + (\Delta_T - \Delta_t) | \mathcal{F}_t] = \mathbb{E}^Q[W_T - W_t | \mathcal{F}_t] + X_t, \tag{7.5}
\]

and so an asset price bubble \( X_t - X^*_t = W_t - \mathbb{E}^Q[W_T | \mathcal{F}_t] \) appears if and only if the local \( Q^* \)-martingale \( W \) is not a true \( Q^* \)-martingale. Unlike in Loewenstein/Willard [39],
a bubble in the bank account is not possible in [30] (nor [31]) due to the more restrictive assumption of NFLVR (instead of NUPBR). As in [39], the link between bubbles and strict local martingales is directly due to the definition (7.5) of the fundamental value. [30] also introduce different types of bubbles (depending on the time-horizon), provide a decomposition of bubbles and discuss the valuation of contingent claims.

The incomplete case in [31] is more challenging. Since $\mathcal{M}_t(W)$ is no longer a singleton, it is not clear which ELMM $Q$ one should use to define a fundamental value as in (7.5). Just picking one $Q$ and using that throughout would be ad hoc and would also just lead back to the complete case results. To address this issue, [31] propose a mechanism where “the market” chooses and sometimes (at random times $\sigma_i$) changes the measure used in (7.5), so that one works with $Q^i \in \mathcal{M}_t(W)$ for times $t$ between $\sigma_i$ and $\sigma_{i+1}$. In effect, this means that one uses a fundamental value of the form

$$X_t^* := \Phi^Q_t(\Delta, X) := E^Q_t[W_T - W_t | \mathcal{F}_t] + X_t,$$

where the measure $Q_t$ used at time $t$ now depends on $t$ as well, and so the analysis of $X - X^*$ becomes more involved. In the same spirit, but in a different setup\footnote{[31] needs a bigger filtration $\mathcal{G}$ to accommodate the $\sigma_i$ (which are independent of $\mathcal{F}$), whereas [1] always stays within $\mathcal{F}$.}, Biagini et al. [1] study the case where $Q_t$ moves smoothly from one $Q$ to another $R$. In both cases, however, we find the actual choice of $Q^i$, or $Q$ and $R$, not fully convincing from an economic perspective. For example, [31] makes the assumptions that there are enough liquidly traded derivatives in the market to determine the ELMM $Q$, and that $Q$ can actually be identified from market prices. This practically leads us back to the complete case studied in [30]. However, we are not aware of any well-established procedures to implement an identification of $Q$ from market prices, and we also find a $Q$ determined from liquid derivative prices conceptually difficult to reconcile with possible violations, due to bubbles, of e.g. put-call parity.

In any case, the resulting difference $X - X^*$ should be called a $Q$-bubble, because there may well be another ELMM $Q'$ under which $W$ is a true martingale (see Example 6.4). So the above notion of a bubble depends in a crucial way on the choice of the risk-neutral measure(s). We prefer an approach which only uses basic assets as given and does not need partly exogenous inputs to define bubbles. Hence, our notion of a bubbly market is more restrictive than the notion of a Q-bubble as in [31]; in fact, if $S = (1, X)$ has a Q-bubble, $S$ generated by $S$ need not be an interesting bubbly market in our sense.

A major part of the analysis in [30, 31] studies issues of valuation in markets with bubbles, and we comment on this below in Section 7.4.Protter [46] also presents ideas to identify a bubble by statistical methods, and gives in Section 11 an overview and discussion of other approaches in the literature. We refer to that instead of repeating it here.

### 7.4 Bubbles and derivative pricing

Bubbles and strict local martingales have come up in mathematical finance with some prominence in the area of option pricing, in particular with relation to violations of put-call parity. Early work on that topic appears in Lewis [37], and this has been taken up in two seminal papers by Cox/Hobson [3] and Heston, Loewenstein and Willard [26].
The setup of [3] is like [30] similar to [39] but a bit more restrictive; they have a model \((1, X, 0)\) without dividends, where \(X \geq 0\) is a continuous semimartingale, and they assume NFLVR and completeness so that they have a unique ELMM \(Q = Q^*\) for \(X (= W\) here). They say that \(X\) has a bubble if it is a strict local \(Q^*\)-martingale; so the definition is the same as in [30]. The main focus of [3] is then on valuation of options in the presence of bubbles, and in particular on violation of put-call parity.

Heston et al. [26] consider a setup \((1, X, 0)\) without dividends, where \(X\) is a one-dimensional local or stochastic volatility model; so they allow in particular incomplete markets. They say that “[a]n asset with a nonnegative price has a “bubble” if there is a self-financing portfolio with pathwise nonnegative wealth that costs less than the asset and replicates the asset’s price at a fixed future date. The bubble’s value is the difference between the asset’s price and the lowest cost replicating strategy” [26, Definition 2.1]. In terms of the discussion in Section 7.1, this means that they use as fundamental value the superreplication price. As already mentioned in Section 7.2, this definition has also appeared in the recent paper by Loewenstein/Willard [40].

The approach of [26, 40] is closely related to ours, but because both only examine the fundamental value at time 0, their notion of a bubble is more restrictive than ours for incomplete markets. Indeed, let \(\mathcal{S}\) be the market generated by \(S = (1, X)\). If there exists an ELMM \(Q\) for \(\mathcal{S}\) (or for \(X\), for that matter), then \(\mathcal{S}\) is dynamically viable (Theorem 4.3); and if \(X\) is a strict local \(Q\)-martingale for each such \(Q\), then \(\mathcal{S}\) is an interesting bubbly market in our sense (Theorem 5.5). Denote by \(\mathcal{M}_e(X)\) the nonempty set of all ELMMs for \(X\). By the classical hedging duality in (7.4), the classical time 0 superreplication price of \(X_t\) is given by \(\sup_{Q \in \mathcal{M}_e(X)} E_Q[X_t], t \in [0, T]\). Now it may happen (see Example 3.7) that \(X_0 = \sup_{Q \in \mathcal{M}_e(X)} E_Q[X_T]\) even though \(X\) is a strict local \(Q\)-martingale for all \(Q \in \mathcal{M}_e(X)\). Then \(\mathcal{S}\) is an interesting bubbly market in our sense, but \(X\) has no bubble in the sense of [26, 40]. For complete markets, as already seen in Section 7.1, the concepts coincide: if there is only one ELMM \(Q^*\) for \(X\), then \(X\) is a strict local \(Q^*\)-martingale if and only if \(X_0 > E_{Q^*}[X_T]\) for some \(t \in [0, T]\). In other words, our approach (but not the one of [26, 40]) allows the possibility of “bubble birth” as in Example 3.7; but this can only happen in an incomplete setting.

The main focus of [26] is on relating the existence of bubbles to multiplicity (nonuniqueness) of solutions to the valuation PDEs of call and put options. But the authors also provide in their specific stochastic volatility framework necessary and sufficient conditions for various bubbles (on the bank account or on the stock). Other papers that study failures of option pricing properties in models with bubbles or strict local martingales are Ekström/Tysk [11] (via PDE techniques), Pal/Protter [43] (via \(h\)-transforms) or Madan/Yor [41] (who connect this to an extension of Itô’s formula), among others.

### 7.5 Bubbles and arbitrage

In recent years, there has been a lot of interest in models which do not satisfy the classic strong absence-of-arbitrage condition NFLVR of Delbaen/Schachermayer [4, 8]. A major motivation has been that a number of empirical observations do not fit well with the stringent properties imposed by NFLVR, with prominent examples being stochastic portfolio theory (Fernholz [14]) or the benchmark approach (Platen/Heath [44]). In most papers
on these subjects, however, the models still satisfy NUPBR, and this brings us close again to the general setup presented in Section 7.1.

One typical question to ask is if or how hedging still works and if or how one could exploit the presence of potential arbitrages. The latter aspect is for example studied in Fernholz et al. [15]. They consider an Itô process model \((1, X, 0)\) where each asset is positive and do not assume NFLVR, but (implicitly) that the market price of risk is \(\mathbb{P}\)-a.s. square-integrable, which implies the existence of a local martingale deflator and hence NUPBR. Their goal is to explicitly construct portfolios which grow more quickly than the market portfolio; this is called “relative arbitrage”. In our terminology, they assume that the market is dynamically viable, but not dynamically efficient. Hence the market portfolio is not maximal, and so it is not surprising that there exist (maximal) strategies which “improve” it. But of course the main contribution of stochastic portfolio theory is to provide explicit formulas for such strategies.

In a similar vein, Ruf [48] discusses “hedging under arbitrage”, i.e., when NFLVR fails but we still have NUPBR. His setup is very similar to Loewenstein/Willard [39, 40], but the market is not assumed to be complete. In a Markovian framework, [48] then uses Feynman–Kac type results to obtain and compute optimal strategies which superreplicate a given contingent claim with minimal initial capital. The link to bubbles here is mainly that by definition, bubbles are identified with assets whose prices follow strict local martingales.

In a quite different direction, one can ask what extra ingredient is needed to avoid having strict local martingales for asset prices, because many pricing anomalies turn out to be consequences of bubbles in that sense. As we have seen in Corollary 5.10, what is needed is a concept of no dominance. This idea goes back to Merton [42] and was formalised mathematically by Jarrow et al. [30, 31] in two different ways; related work was done apparently in parallel by Heston et al. [26]. A detailed study with very clear definitions has then been given by Jarrow and Larsson [29]. They consider a market \((1, X, 0)\), where \(X\) is a \(d\)-dimensional semimartingale for the filtration \(\mathcal{F} = (\mathcal{F}_t)_{t \geq 0}\). They say that “the market \((\mathcal{F}, X)\) is [...] efficient on \([0, T]\) with respect to \(\mathcal{F}\) if there exist a consumption good price index \(\psi\) and an economy \(((P_k)_{k=1}^K, \mathcal{F}, (\epsilon_k)_{k=1}^K, (U_k)_{k=1}^K)\) for which \((\psi, X)\) is an equilibrium price process \(X\) on \([0, T]\)”. Here \(k = 1, \ldots, K\) denote different investors with beliefs \(P_k \approx \mathbb{P}\) (subjective probability measures), endowment streams \(\epsilon_k\) and (time-dependent) utility functions \(U_k\). [29] show that \((\mathcal{F}, X)\) is efficient on \([0, T]\) if and only if \(X\) satisfies NFLVR and no dominance (ND), i.e., the invest-and-keep strategy of each risky asset is maximal, or equivalently if and only if there exists an equivalent (true) martingale measure \(Q \approx \mathbb{P}\) on \(\mathcal{F}_T\). Moreover, they consider the case of different information sets and finally provide examples of efficient and inefficient markets, namely local and stochastic volatility models.

Our definition of dynamic efficiency and dynamic no dominance is directly inspired by [29], and our Corollary 5.10 is a numéraire-independent version of (part of) their key result in Theorem 3.2. In particular, that result justifies our terminology of dynamic efficiency and also motivates our notion of static efficiency. Nevertheless, our definition of dynamic no dominance is (formally) a bit weaker since it only imposes maximality for the market portfolio, not for all individual invest-and-keep strategies in all assets.
7.6 Bubbles and non-bubbles

Despite its numéraire-dependence as discussed in Remark 2.11(a), one important inspiration for many of our concepts has been the work of Delbaen and Schachermayer, especially [6] for numéraire changes and related topics and [7] for maximality. We emphasise again that a direct comparison is delicate because we work with a different notion of admissible strategies. But there is no doubt that F. Delbaen is also well aware of the close connections between maximal elements or strategies, bubbles, and strict local martingales. This is illustrated by a presentation given in June 2012 at the QMF conference in Cairns, Australia. We quote from these slides that “[a] bubble is something that has a price that is too high or for the same amount of money you can get something better” and that “H · S, acceptable, could be called a bubble if the price of \( f = (H \cdot S)_\infty \) is strictly lower than 0”. Delbaen also proposes some ideas to define non-bubbles; however, we have not seen any published work or preprint so far.

Acknowledgements

We gratefully acknowledge financial support by the Swiss Finance Institute (SFI) and the National Centre of Competence in Research “Financial Valuation and Risk Management” (NCCR FINRISK), Project D1 (Mathematical Methods in Financial Risk Management). The NCCR FINRISK is a research instrument of the Swiss National Science Foundation.

We thank the participants at several workshops and conferences for feedback and questions which prompted us to explain our ideas more transparently. Special thanks go to Mark Loewenstein and Freddy Delbaen for useful discussions and insightful comments.

References


