Globally Evolutionarily Stable Portfolio Rules*

Igor V. Evstigneev\textsuperscript{a}, Thorsten Hens\textsuperscript{b} and
Klaus Reiner Schenk-Hoppé\textsuperscript{c}

September 20, 2005

Abstract

The paper examines a dynamic model of a financial market with endogenous asset prices determined by short run equilibrium of supply and demand. Assets pay dividends, that are partially consumed and partially reinvested. The traders use fixed-mix investment strategies (portfolio rules), distributing their wealth between assets in fixed proportions. Our main goal is to identify globally evolutionarily stable strategies, allowing an investor to “survive,” i.e., to accumulate in the long run a positive share of market wealth, regardless of the initial state of the market. It is shown that there is a unique portfolio rule with this property—an analogue of the famous Kelly (1956) rule of “betting one’s beliefs.”

\textit{JEL-Classification:} G11, C61, C62.

\textit{Keywords:} Evolutionary Finance, Wealth Dynamics, Survival and Extinction of Portfolio Rules, Evolutionary Stability, Kelly Rule.

\textsuperscript{*} Financial support by the national center of competence in research “Financial Valuation and Risk Management” is gratefully acknowledged. The national centers in research are managed by the Swiss National Science Foundation on behalf of the federal authorities.

\textsuperscript{a} Economic Studies, School of Social Sciences, University of Manchester, Oxford Road, Manchester M13 9PL, United Kingdom. E-mail: igor.evstigneev@manchester.ac.uk.

\textsuperscript{b} Institute for Empirical Research in Economics, University of Zurich, Blümlisalpstrasse 10, 8006 Zürich, Switzerland and Norwegian School of Economics and Business Administration, Helleveien 30, N-5045 Bergen, Norway. E-mail: thens@iew.unizh.ch.

\textsuperscript{c} Leeds University Business School and School of Mathematics, University of Leeds, Leeds LS2 9JT, United Kingdom. E-mail: K.R.Schenk-Hoppe@leeds.ac.uk.
1 Introduction

Price changes and dividend payments of stocks induce wealth dynamics among investors using different investment strategies (portfolio rules) in financial markets. These dynamics act as a natural selection force among the portfolio rules: some prove to be successful and “survive,” the others fail and "become extinct.” The purpose of the present paper is to investigate financial dynamics from this evolutionary perspective with the view of identifying evolutionarily stable (surviving) investment strategies.

Evolutionary ideas have a long history in the social sciences going back to Malthus, who played an inspirational role for Darwin (for a review of the subject see, e.g., Hodgeson [16]). A more recent stage of development of these ideas began in the 1950s with the publications of Akhian [1], Penrose [22] and others. A powerful momentum to work in this area was given by the interdisciplinary research conducted in the 1980s and 1990s under the auspices of the Santa Fe Institute in New Mexico, USA, where researchers of different backgrounds—economists, mathematicians, physicists and biologists—combined their efforts to study evolutionary dynamics in biology, economics and finance; see, e.g., Arthur, Holland, LeBaron, Palmer and Taylor [4], Farmer and Lo [14], LeBaron, Arthur and Palmer [18], Blume and Easley [6], and Blume and Durlauf [5].

While inspired by the above studies, especially by the pioneering work of Blume and Easley [6], our approach to evolutionary finance is different from theirs both in the modeling frameworks and in the specific problems analyzed. In particular, we deal with models based on random dynamical systems, rather than on the conventional general equilibrium settings where agents maximize discounted sums of expected utilities. The emphasis is on finding explicit formulas for evolutionarily stable portfolio rules with the view of making the theory closer to practical applications. In contrast to a number of the papers mentioned above, we use mathematical modeling, rather than computer simulations, to justify our conclusions.

We consider a dynamic stochastic model of a financial market in which there are $I$ investors (traders) and $K$ traded assets (securities). Asset supply is constant over time. Each trader chooses a strategy prescribing to distribute, at the beginning of each time period $t=1,2,...,$ his/her investment budget between the securities according to fixed proportions. Assets pay dividends, that are random and depend on a discrete-time stochastic process of exogenous “states of the world.”

The prices of the securities at each date are derived endogenously from the equilibrium condition: aggregate market demand of each asset is equal to its supply. Each investor’s individual demand depends on his/her bud-
get and investment strategy. The latter is fixed (the investment proportions are constant over time). The former depends on time and random factors. The investment budget has two sources: the dividends paid by the assets and capital gains. These two sources form investor's wealth, which is partially consumed and partially reinvested at each time period. The investment/consumption ratio is fixed, and it is supposed to be the same for all the traders.

We note that the class of fixed-mix, or constant proportions, strategies we consider in this work is quite common in financial theory and practice; see, e.g., Perold and Sharpe [23], Mulvey and Ziemba [19], Browne [8] and Dempster [10, 11].

The strategy profile of the investors determines the "ecology" of the market and its random dynamics over time. In the evolutionary perspective, survival or extinction of investment strategies is governed by the long-run behavior of the relative wealth of the investors, which depends on the combination of the strategies chosen. A portfolio rule (or an investor using it) is said to survive if it accumulates in the limit a positive fraction of total market wealth. It is said to become extinct if the share of market wealth corresponding to it tends to zero.

An investment strategy, or a portfolio rule, is called evolutionarily stable if the following condition holds. If a group of investors uses this rule, while all the others use different ones, those and only those investors survive who belong to the former group. If this condition holds regardless of the initial state of the market, the investment strategy is called globally evolutionarily stable. If it holds under the additional assumption that the group of investors using other portfolio rules (distinct from the one we consider) possesses a sufficiently small initial share of market wealth, then the above property of stability is termed local.

We prove that among all fixed-mix investment strategies, the only globally evolutionarily stable portfolio rule is to invest according to the proportions of the expected dividends. This recipe is similar to the well-known Kelly's principle of "betting one's beliefs." The present paper contributes to that field of studies which originated from the pioneering work of Shannon\(^1\) and Kelly [17]—see Breiman [7], Thorp [24], Algoet and Cover [2], Hakansson and Ziemba [15] and references therein. Most of the previous work was concerned with models where asset prices were given exogenously, or where

\(^{1}\)Although Claude Shannon—the famous founder of the mathematical theory of information—did not publish on investment-related issues, his ideas, expressed in his lectures on investment problems, should apparently be regarded as the initial source of that strand of literature which we cite here. For the history of these ideas and the related discussion see Cover [9].
the analysis was based on a reduction to such models [6]. Our aim is to obtain analogous results in a dynamic equilibrium setting, with endogenous prices. Intermediate steps towards this aim were made in the previous papers [3] and [12]. Those papers dealt with a special case of “short-lived” assets. Here, we extend the results to a model with long-lived, dividend-paying assets and thus achieve the long-sought goal of providing a natural and general framework for this class of results.

The structure of the paper is as follows. Section 2 provides a rigorous description of the model, a brief outline of which was given above. In Section 3, we formulate and discuss the main result. Sections 4–6 develop methods needed for the analysis of the model under consideration. Section 7 completes the proof of the main theorem based on the auxiliary results of the foregoing part of the paper. The Appendix contains a technical lemma used in this work.

2 The model

There are \( I \geq 2 \) investors (traders) acting in a market where \( K \geq 2 \) different risky assets, or securities, are traded. Total amount of each security \( k = 1, ..., K \) in the market is constant in time and normalized to 1. At each date \( t = 1, 2, ..., \), the assets \( k = 1, ..., K \) pay dividends \( D_k(s_t) \), where \( s_t \) is the state of the world at date \( t \). The states of the world \( s_1, s_2, ... \) form a sequence of independent identically distributed random elements with values in a set \( S \). The set \( S \) is supposed to be finite, and, for each \( s \in S \), the probability that \( s_t = s \) is strictly positive. The functions \( D_k(s) \), \( k = 1, ..., K \), are non-negative, their sum with respect to \( k \) is strictly positive for each \( s \):

\[
\sum_{k=1}^{K} D_k(s) > 0, \ s \in S, \quad (1)
\]

and

\[
ED_k(s_t) > 0, \ k = 1, ..., K, \quad (2)
\]

where “\( E \)” stands for the expectation with respect to the underlying probability measure \( P \).

Each investor \( i \) chooses an investment strategy (portfolio rule) characterized by a non-negative vector

\[
\lambda^i = (\lambda^i_1, ..., \lambda^i_K)
\]
such that
\[ \lambda_1^i + \ldots + \lambda_K^i = 1. \]

The set of such vectors—the unit simplex in the $K$-dimensional space $\mathbb{R}^K$—will be denoted by $\Delta^K$. The numbers $\lambda_k^i$ indicate the proportions according to which investor $i$ distributes his/her budget between the assets $k = 1, \ldots, K$. These proportions remain the same over time, so that we deal here with simple, or fixed-mix, investment strategies. Throughout this paper we will consider only those portfolio rules $(\lambda_1^i, \ldots, \lambda_K^i)$ which are completely mixed, i.e., $\lambda_k^i > 0$ for each $k = 1, \ldots, K$.

Asset prices $p_{t,1}, \ldots, p_{t,K}$ at each time period $t$ are determined by the equilibrium condition: demand equals supply. Total amount invested (by all the traders) in asset $k$ at date $t$ equals

\[ \langle \lambda_k, b_t \rangle = \sum_{i=1}^{I} \lambda_k^i b_t^i, \]

where $b_t^i$ is the trader $i$’s investment budget at date $t$ and $b_t = (b_t^1, \ldots, b_t^I)$. Equality

\[ \frac{\langle \lambda_k, b_t \rangle}{p_{t,k}} = 1, \]

expressing the fact that demand of asset $k$ is equal to its supply, gives the formula for the equilibrium price of asset $k$ at date $t$:

\[ p_{t,k} = \langle \lambda_k, b_t \rangle. \tag{3} \]

Investor $i$ with budget $b_t^i$, distributing it between the assets according to the proportions $\lambda_1^i, \ldots, \lambda_K^i$, purchases

\[ x_{t,k}^i = \frac{\lambda_k^i b_t^i}{p_{t,k}} = \frac{\lambda_k^i b_t^i}{\langle \lambda_k, b_t \rangle} \tag{4} \]

units of asset $k$ at date $t$. The number $x_{t,k}^i$ is equal to the amount of money $\lambda_k^i b_t^i$ investor $i$ spends for purchasing security $k$ divided by the price $p_{t,k} = \langle \lambda_k, b_t \rangle$ of this security. Thus the vector

\[ x_t^i = (x_{t,1}^i, \ldots, x_{t,K}^i) \]

is the portfolio of investor $i$ at date $t$. (Here, positions $x_{t,k}^i$ of the portfolio $x_t^i$ are measured in terms of “physical units” of assets.)
At the beginning of the next time period \( t+1 \), investor \( i \) obtains *dividends* from the portfolio \( x_t^i \) resulting in the amount

\[
\sum_{k=1}^{K} D_k(s_{t+1})x_{t,k}^i = \sum_{k=1}^{K} D_k(s_{t+1}) \frac{\lambda_k b_t^i}{\lambda_k b_t^i}.
\]

Trader \( i \)'s budget at date \( t+1 \) used for purchasing assets is as follows:

\[
b_{t+1}^i = \rho w_{t+1}^i,
\]

where

\[
w_{t+1}^i = \sum_{k=1}^{K} p_{t+1,k} x_{t,k}^i + \sum_{k=1}^{K} D_k(s_{t+1}) x_{t,k}^i.
\]

The first sum in (6) expresses the value of the portfolio \( x_t^i \) in terms of the prices \( p_{t+1,k} \) prevailing at date \( t+1 \). The second sum is the amount of dividends obtained from the portfolio \( x_t^i \). Trader \( i \)'s *wealth* \( w_{t+1}^i \) is divided between investment and consumption in the proportions \( \rho \) and \( 1-\rho \). The number \( \rho \in (0,1) \) is given: \( 1-\rho \) is the *consumption rate*—the same for all the investors. The amount \( b_{t+1}^i = \rho w_{t+1}^i \) is invested into risky assets, while the sum \( (1-\rho)w_{t+1}^i \) is consumed.

We suppose that the consumption rate is the same for all the market traders because we are mainly interested in comparing the long-run performance of investment strategies. This can be done only for a group of traders having the same consumption rate. Otherwise, a seemingly lower performance of a strategy may be simply due to a higher consumption rate of the investor.

Denote by \( W_t \) the *total market wealth* at date \( t \):

\[
W_t = \sum_{i=1}^{I} w_t^i.
\]

Consider the *relative wealth* \( r_t^i \) of investor \( i \), \( i = 1, \ldots, I \), at date \( t \):

\[
r_t^i = \frac{w_t^i}{W_t}.
\]

We are interested in the long-run behavior of the relative wealth, i.e. in the asymptotic properties of the sequence of vectors \( r_t = (r_t^1, \ldots, r_t^I) \) as \( t \to \infty \). To analyze these properties, we will derive equations allowing to compute
the vector \( r_{t+1} \) based on the knowledge of the vector \( r_t \) and the state of the world \( s_{t+1} \) realized at date \( t+1 \). By using the formulas (4) and

\[
p_{t+1,k} = \langle \lambda_k, b_{t+1} \rangle = \rho(\lambda_k, w_{t+1}) \left[ w_t = (w_t^1, ..., w_t^I) \right],
\]

(see (3)), we substitute the values of \( x_{t,k}^i \) and \( p_{t,k} \) into (6), which yields

\[
w_t^i = \sum_{k=1}^{K} \left[ \rho(\lambda_k, w_{t+1}) + D_k(s_{t+1}) \right] \frac{\lambda_k^i w_t^i}{\langle \lambda_k, w_t \rangle}.
\]  \hspace{1cm} (7)

By summing up these equations over \( i = 1, ..., I \), we obtain

\[
W_{t+1} = \sum_{k=1}^{K} \left[ \rho(\lambda_k, w_{t+1}) + D_k(s_{t+1}) \right] \frac{\sum_{i=1}^{I} \lambda_k^i w_t^i}{\langle \lambda_k, w_t \rangle} = \rho W_{t+1} + \sum_{k=1}^{K} D_k(s_{t+1}).
\]

This leads to the formula

\[
W_{t+1} = \frac{D(s_{t+1})}{1 - \rho},
\]  \hspace{1cm} (8)

where

\[
D(s_{t+1}) = \sum_{k=1}^{K} D_k(s_{t+1}) \ (> 0)
\]

is the sum of the dividends of all the assets. Dividing both sides of equation (7) by \( W_{t+1} \) and using formula (8), we find

\[
r_t^i = \sum_{k=1}^{K} \left[ \rho(\lambda_k, r_{t+1}) + (1 - \rho) \frac{D_k(s_{t+1})}{D(s_{t+1})} \right] \frac{\lambda_k^i r_t^i}{\langle \lambda_k, r_t \rangle}.
\]

Thus we arrive at the system of equations:

\[
r_t^i = \sum_{k=1}^{K} \left[ \rho(\lambda_k, r_{t+1}) + (1 - \rho) R_k(s_{t+1}) \right] \frac{\lambda_k^i r_t^i}{\langle \lambda_k, r_t \rangle}, \quad i = 1, ..., I,
\]  \hspace{1cm} (9)

where

\[
R_k(s_{t+1}) = \frac{D_k(s_{t+1})}{D(s_{t+1})}, \quad k = 1, ..., K,
\]

are the relative dividends of the assets \( k = 1, ..., K \). It can be shown (see Section 4 below) that, for any \( r_t = (r_t^1, ..., r_t^I) \in \Delta^I \) and any \( s_{t+1} \in S \), this
system of equations has a unique solution \( r_{t+1} \geq 0 \). We have \( r_{t+1} \in \Delta^I \), which can be verified by summing up equations (9) and using the fact that

\[
\sum_{k=1}^{K} R_k(s) = 1, \ s \in S.
\]

We will denote the solution \( r_{t+1} \) to system (9) (as a function of \( s_{t+1} \) and \( r_t \)) by \( F(s_{t+1}, r_t) \). The mapping \( F(s_{t+1}, \cdot) \) transforms \( \Delta^I \) into \( \Delta^I \). Thus we deal here with a random dynamical system

\[
r_{t+1} = F(s_{t+1}, r_t)
\]

(10)
on the unit simplex \( \Delta^I \). We will assume that a strictly positive non-random vector \( r_0 \in \Delta^I \) is fixed. Starting from this initial state, we can generate a path (trajectory)

\[
r_0, r_1(s^1), r_2(s^2), \ldots
\]
of the random system (10) according to the formula

\[
r_{t+1}(s^{t+1}) = F(s_{t+1}, r_t(s^t)), \ t = 0, 1, \ldots,
\]

where

\[
s^t = (s_1, \ldots, s_t).
\]

(If \( t = 0 \), we formally write \( r_0 = r_0(s^0) \) having in mind that \( r_0 \) is a constant.)

**Remark.** The model we have described was proposed in [13]. Its presentation in this paper is slightly different from that in [13]. (In particular, we here write \( \rho \) in place of \( 1 - \lambda_0 \) and \( \lambda_k^i \) instead of \( \lambda_k^i/(1 - \lambda_0) \) in [13].) In the limit as \( \rho \to 0 \), the model reduces to the one studied in [12]. In particular, if \( \rho = 0 \), the random dynamical system described by equations (9) coincides with that examined in [12].

### 3 The main result

We examine the dynamics of the relative wealth \( r^i_t \), governed by equations (9), from an evolutionary perspective. We are interested in questions of “survival and extinction” of portfolio rules. We say that a portfolio rule \( \lambda^i = (\lambda^i_1, \ldots, \lambda^i_K) \) (or investor \( i \) using it) *survives* with probability one in the market selection process (9) if, for the relative wealth \( r^i_t \) of investor \( i \), we have \( \lim_{t \to \infty} r^i_t > 0 \) almost surely. We say that \( \lambda^i \) *becomes extinct* with probability one if \( \lim_{t \to \infty} r^i_t = 0 \) almost surely.

A central role in this work is played by the following definition.
Definition 1 A portfolio rule $\lambda = (\lambda_1, \ldots, \lambda_K)$ is called globally evolutionarily stable if the following condition holds. Suppose, in a group of investors $i = 1, 2, \ldots, J$ ($1 \leq J < I$), all use the portfolio rule $\lambda$, while all the others, $i = J + 1, \ldots, I$ use portfolio rules $\lambda^i$ distinct from $\lambda$. Then those investors who belong to the former group ($i = 1, \ldots, J$) survive with probability one, whereas those who belong to the latter ($i = J + 1, \ldots, I$) become extinct with probability one.

In the above definition, it is supposed that the initial state $r_0$ in the market selection process governed by equations (9) is any strictly positive vector $r_0 \in \Delta^I$. This is reflected in the term “global evolutionary stability.” An analogous local concept (cf. [13]) is defined similarly, but in the definition of local evolutionary stability, the initial market share $r_0^J + \ldots + r_0^I$ of the group of investors who use strategies $\lambda^i$ distinct from $\lambda$ is supposed to be small enough.

Our main goal is to identify that portfolio rule which is globally evolutionarily stable. Clearly, if it exists it must be unique. Indeed if there are two such rules, $\lambda \neq \lambda'$, we can divide the population of investors into two groups assuming that the first uses $\lambda$ and the second $\lambda'$. Then, according to the definition of global evolutionary stability, both groups must become extinct with probability one, which is impossible since the sum of the relative wealth of all the investors is equal to one.

Define

$$\lambda^* = (\lambda_1^*, \ldots, \lambda_K^*), \quad \lambda_k^* = ER_k(s), \quad k = 1, \ldots, K,$$

so that $\lambda_1^*, \ldots, \lambda_K^*$ are the expected relative dividends of assets $k = 1, \ldots, K$. The portfolio rule (investment strategy) $\lambda^*$ is called the Kelly portfolio rule.

It prescribes to invest in accordance with the principle of “betting one’s beliefs,” as formulated in the pioneering paper by Kelly [17], for further studies in this direction see Breiman [7], Thorp [24], Algoet and Cover [2] and Hakansson and Ziemba [15].

Recall that, according to a convention made in Section 2, we consider in this paper only completely mixed portfolio rules. Therefore the vectors $\lambda$ and $\lambda^i$ involved in Definition 1 are supposed to be strictly positive. The Kelly rule is completely mixed by virtue of assumptions (1) and (2).

Throughout the paper, we will assume that the functions $R_1(s), \ldots, R_K(s)$ are linearly independent (there are no redundant assets).

The main result of this paper is as follows.

Theorem 1 The Kelly rule is globally evolutionarily stable.
In order to prove this theorem we have to consider a group of investors \( i = 1, ..., J \) using the portfolio rule \( \lambda^* \), assume that all the other investors \( i = J+1, ..., I \) use portfolio rules \( \lambda^i \neq \lambda^* \) and show that the former group survives, while the latter becomes extinct. In general, \( J \) should be any number between \( 1 \leq J < I \). We note, however, that it is sufficient to prove the theorem assuming that \( J = 1 \), in which case the result reduces to the assertion that \( r^1_t \rightarrow 1 \) almost surely. To perform the reduction of the case \( J > 1 \) to the case \( J = 1 \), we “aggregate” the group of investors \( i = 1, 2, ..., J \) into one by setting

\[
\bar{r}^1_t = r^1_t + ... + r^J_t.
\]

By adding up equations (9) over \( i = 1, ..., J \), we obtain:

\[
\bar{r}^1_{t+1} = \sum_{k=1}^{K} [\rho(\lambda_k, r_{t+1}) + (1 - \rho) R_k(s_{t+1})] \frac{\lambda_k^* r^1_t}{(\lambda_k, r_t)},
\]

where

\[
\langle \lambda_k, r \rangle = \lambda_k^* r^1 + \sum_{i=J+1}^{I} \lambda_k^i r^i.
\]

Thus the original model reduces to the analogous one in which there are \( I - J + 1 \) investors \((i = 1, J+1, ..., I)\) so that investor 1 uses the Kelly strategy \( \lambda^* \) and all the others, \( i = J+1, ..., I \), use strategies distinct from \( \lambda^* \). If we have proved Theorem 1 in the special case \( J = 1 \), we know that \( r^i_t \rightarrow 0 \) almost surely for all \( i = J+1, ..., I \). Consequently, \( \bar{r}^1_t \rightarrow 1 \), which means that the group of investors \( i = 1, ..., I \) (which we treat as a single, “aggregate,” investor) accumulates in the limit all market wealth. It remains to observe that in the original model, the proportions between the relative wealth of investors \( i, j \) who belong to the group \( 1, ..., J \) using the Kelly rule do not change in time. This is so because for all such investors, the growth rates of their relative wealth are the same:

\[
\frac{r^i_{t+1}}{r^i_t} = \sum_{k=1}^{K} [\rho(\lambda_k, r_{t+1}) + (1 - \rho) R_k(s_{t+1})] \frac{\lambda_k^*}{\langle \lambda_k, r_t \rangle}, \quad i = 1, ..., J.
\]

Consequently,

\[
\frac{r^i_{t+1}}{r^i_t} = \frac{r^j_{t+1}}{r^j_t}, \quad i, j = 1, ..., J,
\]

\[10\]
and so
\[
\frac{r^i_{t+1}}{r^i_{t+1}} = \frac{r^i_t}{r^i_t}, \quad i = 1, \ldots, J.
\]
Therefore \( r^i_t = \beta^i r^i_t \) (\( i = 1, \ldots, J \)) for all \( t \), where \( \beta^i = r^i_{0}/r^i_0 \) is a strictly positive constant. Since
\[
\sum_{i=1}^{J} r^i_t = (\sum_{i=1}^{J} \beta^i) r^1_t \to 1 \quad \text{(a.s.)},
\]
we obtain that \( r^i_t \to \beta^i (\sum_{i=1}^{J} \beta^i)^{-1} > 0 \) (a.s.) for all \( i = 1, \ldots, J \), which means that all the “Kelly investors” \( i = 1, \ldots, J \) survive.

Thus in order to prove Theorem 1 it is sufficient to establish the following fact: if investor 1 uses the Kelly rule, while all the others use strategies distinct from the Kelly rule, investor 1 is almost surely the single survivor in the market selection process. We will prove this assertion in Section 7 based on a number of auxiliary results which will be obtained in Sections 4-6. These results provide methods needed for the analysis of the model at hand, and some of them (especially those in Section 5) are of independent interest.

4 The mapping defining the random dynamical system

Let \( \rho \) be a number satisfying \( 0 \leq \rho < 1 \). For each \( s \in S \), consider the mapping

\[
F(s, r) = (F^1(s, r), \ldots, F^I(s, r))
\]

of the unit simplex \( \Delta^I \) into itself defined by

\[
F^i(s, r) = \sum_{k=1}^{K} [\rho(\lambda_k, F(s, r)) + (1 - \rho)R_k(s)] \frac{\lambda_k r^i}{\langle \lambda_k, r \rangle}, \quad i = 1, \ldots, I.
\]  

(11)

The fact that the mapping under consideration is well-defined is established in Proposition 1 below. Fix some element \( s \) of the state space \( S \) and a vector \( r = (r^1, \ldots, r^I) \in \Delta^I \). Consider the affine operator \( B : \mathbb{R}^I_+ \to \mathbb{R}^I_+ \) transforming a vector \( x = (x^1, \ldots, x^I) \in \mathbb{R}^I_+ \) into the vector \( y = (y^1, \ldots, y^I) \in \mathbb{R}^I_+ \) defined by

\[
y^i = \sum_{k=1}^{K} [\rho(\lambda_k, x) + (1 - \rho)R_k] \frac{\lambda_k r^i}{\langle \lambda_k, r \rangle},
\]

where \( R_k = R_k(s) \).
Proposition 1 The operator $B$ possesses a unique fixed point in $\mathbb{R}^I_+$. This point belongs to the unit simplex $\Delta^I$.

Proof. Consider any $x, \bar{x} \in \mathbb{R}^I_+$ and put $y = B(x)$, $\bar{y} = B(\bar{x})$. We have

$$|y - \bar{y}| = \sum_{i=1}^{I} |y^i - \bar{y}^i| = \rho \sum_{i=1}^{I} \sum_{k=1}^{K} \langle \lambda_k, x - \bar{x} \rangle \frac{\lambda_k r^i}{\langle \lambda_k, r \rangle} \leq$$

$$\rho \sum_{k=1}^{K} \sum_{i=1}^{I} |\langle \lambda_k, x - \bar{x} \rangle| \frac{\lambda_k r^i}{\langle \lambda_k, r \rangle} = \rho \sum_{k=1}^{K} |\langle \lambda_k, x - \bar{x} \rangle| \leq$$

$$\rho \sum_{k=1}^{K} \sum_{j=1}^{I} \lambda_k^j |x^j - \bar{x}^j| = \rho \sum_{j=1}^{I} |x^j - \bar{x}^j| = \rho |x - \bar{x}|.$$

Thus the operator $B : \mathbb{R}^I_+ \to \mathbb{R}^I_+$ is contracting and hence it contains a unique fixed point $x \in \mathbb{R}^I_+$. To show that $x \in \Delta^I$ we sum up the equations

$$x^i = \sum_{k=1}^{K} [\rho \langle \lambda_k, x \rangle + (1 - \rho) R_k] \frac{\lambda_k r^i}{\langle \lambda_k, r \rangle}$$

over $i = 1, ..., I$ and obtain

$$|x| = \sum_{k=1}^{K} [\rho \langle \lambda_k, x \rangle + (1 - \rho) R_k] \frac{\sum_{i=1}^{I} \lambda_k r^i}{\langle \lambda_k, r \rangle} = \sum_{k=1}^{K} [\rho \langle \lambda_k, x \rangle + (1 - \rho) R_k] =$$

$$\rho |x| + (1 - \rho),$$

which yields $|x| = 1$. \hfill \Box

Proposition 2 We have

$$\sum_{i=1}^{I} |F^i(s, r) - F^i(s, \bar{r})| \leq \frac{1}{(1 - \rho)} \sum_{i=1}^{I} \sum_{k=1}^{K} \frac{\lambda_k r^i}{\langle \lambda_k, r \rangle} - \frac{\lambda_k \bar{r}^i}{\langle \lambda_k, \bar{r} \rangle}, \quad r, \bar{r} \in \Delta^I. \quad (12)$$

12
It follows from (12) and the inequalities

\[
\langle \lambda_k, r \rangle = \sum_{i=1}^{I} \lambda_k^i r^i > 0, \quad \langle \lambda_k, \tilde{r} \rangle = \sum_{i=1}^{I} \lambda_k^i \tilde{r}^i > 0,
\]

(holding because \( \lambda_k^i > 0 \)) that the mapping \( F(s, r) \) is continuous in \( r \in \Delta^I \).

**Proof of Proposition 2.** For any \( r, \tilde{r} \in \Delta^I \) and \( i = 1, \ldots, I \), we have

\[
|F^i(s, r) - F^i(s, \tilde{r})| = \left| \sum_{k=1}^{K} \left[ \rho(\lambda_k, F(s, r)) + (1 - \rho)R_k(s) \right] \frac{\lambda_k^i r^i}{\langle \lambda_k, r \rangle} - \sum_{k=1}^{K} \left[ \rho(\lambda_k, F(s, \tilde{r})) + (1 - \rho)R_k(s) \right] \frac{\lambda_k^i \tilde{r}^i}{\langle \lambda_k, \tilde{r} \rangle} \right|
\]

\[
\sum_{k=1}^{K} \left[ \rho(\lambda_k, F(s, \tilde{r})) + (1 - \rho)R_k(s) \right] \frac{\lambda_k^i \tilde{r}^i}{\langle \lambda_k, \tilde{r} \rangle} \leq \rho \sum_{k=1}^{K} \left| \langle \lambda_k, F(s, r) \rangle \left[ \frac{\lambda_k^i r^i}{\langle \lambda_k, r \rangle} - \frac{\lambda_k^i \tilde{r}^i}{\langle \lambda_k, \tilde{r} \rangle} \right] \right|
\]

\[
(1 - \rho) \sum_{k=1}^{K} \left| \frac{\lambda_k^i r^i}{\langle \lambda_k, r \rangle} - \frac{\lambda_k^i \tilde{r}^i}{\langle \lambda_k, \tilde{r} \rangle} \right| \leq \rho \sum_{k=1}^{K} \left| \langle \lambda_k, F(s, r) \rangle - \langle \lambda_k, F(s, \tilde{r}) \rangle \right| \frac{\lambda_k^i \tilde{r}^i}{\langle \lambda_k, \tilde{r} \rangle} + \rho \sum_{k=1}^{K} \left| \langle \lambda_k, F(s, r) \rangle - \langle \lambda_k, F(s, \tilde{r}) \rangle \right| \frac{\lambda_k^i r^i}{\langle \lambda_k, r \rangle} + (1 - \rho) \sum_{k=1}^{K} \left| \frac{\lambda_k^i r^i}{\langle \lambda_k, r \rangle} - \frac{\lambda_k^i \tilde{r}^i}{\langle \lambda_k, \tilde{r} \rangle} \right|
\]

\[
\rho \sum_{k=1}^{K} \left| \langle \lambda_k, F(s, r) \rangle - \langle \lambda_k, F(s, \tilde{r}) \rangle \right| \frac{\lambda_k^i \tilde{r}^i}{\langle \lambda_k, \tilde{r} \rangle} + \rho \sum_{k=1}^{K} \left| \langle \lambda_k, F(s, r) \rangle - \langle \lambda_k, F(s, \tilde{r}) \rangle \right| \frac{\lambda_k^i r^i}{\langle \lambda_k, r \rangle} + (1 - \rho) \sum_{k=1}^{K} \left| \frac{\lambda_k^i r^i}{\langle \lambda_k, r \rangle} - \frac{\lambda_k^i \tilde{r}^i}{\langle \lambda_k, \tilde{r} \rangle} \right|
\]

\[
\rho \sum_{k=1}^{K} \left| \langle \lambda_k, F(s, r) \rangle - \langle \lambda_k, F(s, \tilde{r}) \rangle \right| \frac{\lambda_k^i \tilde{r}^i}{\langle \lambda_k, \tilde{r} \rangle} + \sum_{k=1}^{K} \left| \frac{\lambda_k^i r^i}{\langle \lambda_k, r \rangle} - \frac{\lambda_k^i \tilde{r}^i}{\langle \lambda_k, \tilde{r} \rangle} \right|
\]
By summing up these inequalities over \( i = 1, \ldots, I \), we obtain

\[
\sum_{i=1}^{I} |F^i(s, r) - F^i(s, \tilde{r})| \leq \\
\rho \sum_{k=1}^{K} |\langle \lambda_k, F(s, r) \rangle - \langle \lambda_k, F(s, \tilde{r}) \rangle| + \rho \sum_{i=1}^{I} \sum_{k=1}^{K} \left| \frac{\lambda^i_{k,r^i}}{\langle \lambda_k, r \rangle} - \frac{\lambda^i_{k,\tilde{r}^i}}{\langle \lambda_k, \tilde{r} \rangle} \right| \leq \\
\rho \sum_{i=1}^{I} |F^i(s, r) - F^i(s, \tilde{r})| + \sum_{i=1}^{I} \sum_{k=1}^{K} \left| \frac{\lambda^i_{k,r^i}}{\langle \lambda_k, r \rangle} - \frac{\lambda^i_{k,\tilde{r}^i}}{\langle \lambda_k, \tilde{r} \rangle} \right|,
\]

which yields (12). \( \square \)

For each \( s \in S \) and \( r = (r^1, \ldots, r^I) \in \Delta^I \), define

\[
g^i(s, r) = \sum_{k=1}^{K} [\rho(\lambda_k, F(s, r)) + (1 - \rho)R_k(s)] \frac{\lambda^i_k}{\langle \lambda_k, r \rangle}, \quad i = 1, \ldots, I. \tag{13}
\]

It follows from (11) that if \( r^i > 0 \), then

\[
g^i(s, r) = \frac{F^i(s, r)}{r^i}
\]

so that \( g^i(s, r) \) is the growth rate of the \( i \)th coordinate under the mapping \( F \). Define

\[
\mu_* = \min_{i,k} \lambda^i_k, \quad \mu^* = \max_{i,k} \lambda^i_k, \quad H = \mu^*/\mu_*.
\]

The proposition below shows that the growth rate is uniformly bounded away from zero and infinity.

**Proposition 3** For each \( r \in \Delta^I \) and each \( i = 1, \ldots, I \), we have

\[
H^{-1} \leq g^i(s, r) \leq H, \quad s \in S. \tag{14}
\]

The function \( g^i(s, r) \) is continuous in \( r \in \Delta^I \).

**Proof.** Since \( \mu_* \leq \langle \lambda_k, r \rangle \leq \mu^* \), we obtain

\[
H^{-1} = \frac{\mu_*}{\mu^*} \leq \frac{\lambda^i_k}{\langle \lambda_k, r \rangle} \leq \frac{\mu^*}{\mu_*} = H,
\]

which yields (14) because

\[
\sum_{k=1}^{K} \rho(\lambda_k, F(s, r)) + (1 - \rho)R_k(s) = \rho \sum_{k=1}^{K} \langle \lambda_k, F(s, r) \rangle + (1 - \rho) \sum_{k=1}^{K} R_k(s) = 1.
\]

The function \( g^i(s, r) \) is continuous in \( r \in \Delta^I \), because \( F(s, r) \) is continuous in \( r \) and \( \langle \lambda_k, r \rangle \geq \mu_* > 0 \) (see (13)). \( \square \)
5 Return on the Kelly portfolio

Define

\[ f(s, r) = \sum_{k=1}^{K} \left[ \rho(\lambda_k, F(s, r)) + (1 - \rho) R_k(s) \right] \lambda_k^* \]

where \( \lambda_k^* = ER_k(s), \ k = 1, ..., K \). Suppose, at date \( t \), the relative wealth of the investors \( i = 1, ..., I \) are given by the vector \( r = (r^1, ..., r^I) \in \Delta^I \). Then the (relative) asset prices at dates \( t \) and \( t + 1 \) are

\[ p_k = \langle \lambda_k, r \rangle, \quad q_k(s) = \langle \lambda_k, F(s, r) \rangle, \]

provided the state of the world realized at date \( t + 1 \) is \( s \). An investor’s portfolio in which unit wealth is distributed between the assets according to the proportions \( \lambda_k^* \), \( k = 1, ..., K \), is called the Kelly portfolio. The (gross) return on this portfolio, taking into account the dividends and consumption, is given by the function \( f(s, r) \) defined by (15), which can be written as

\[ f(s, r) = \sum_{k=1}^{K} \left[ \rho q_k(s) + (1 - \rho) R_k(s) \right] \frac{\lambda_k^*}{p_k}. \]

If one of the investors \( i = 1, ..., I \), say investor 1, employs the investment strategy \( \lambda^* = (\lambda_1^*, ..., \lambda_K^*) \) (i.e., \( \lambda_k^* = \lambda_k^*, \ k = 1, ..., K \)), then the growth rate of his/her market share is equal to \( f(s, r) \):

\[ g^1(s, r) = f(s, r) \]

(see (13) and (15)).

An important result on which the analysis of the model at hand is based is contained in the following theorem.

**Theorem 2** For each \( r \in \Delta^I \), we have

\[ E \ln f(s, r) \geq 0. \] \hfill (17)

This inequality is strict if and only if

\[ \langle \lambda_k, r \rangle \neq \lambda_k^* \text{ for at least one } k = 1, ..., K. \] \hfill (18)

This result means that the expected logarithmic return on the Kelly portfolio \((\lambda_1^*, ..., \lambda_K^*)\) is non-negative. It is strictly positive if and only if \((\lambda_1^*, ..., \lambda_K^*)\) does not coincide with the market portfolio \((p_1, ..., p_K)\). Recall
that the total amount of each asset is normalized to 1, so that the total wealth invested into asset \( k \) is \( p_k = \langle \lambda_k, r \rangle \). We emphasize that, in Theorem 2, it is not assumed that any of the investors \( i = 1, \ldots, I \) uses the Kelly strategy. The result is applicable without this assumption.

**Proof of Theorem 2. 1st step.** Multiplying both sides of (11) by \( \lambda_m^i \) and summing up over \( i = 1, \ldots, I \), we get

\[
\langle \lambda_m, F(s, r) \rangle = \sum_{k=1}^{K} \left[ \rho \langle \lambda_k, F(s, r) \rangle + (1 - \rho) R_k(s) \right] \frac{\sum_{i=1}^{I} \lambda_m^i \lambda_k^i r^i}{\langle \lambda_k, r \rangle} \quad (19)
\]

\((m = 1, \ldots, K)\). By using the notation introduced in (16), equations (19) and inequality (17) can be written as

\[
q_m(s) = \sum_{k=1}^{K} \left[ \rho q_k(s) + (1 - \rho) R_k(s) \right] \frac{\sum_{i=1}^{I} \lambda_m^i \lambda_k^i r^i}{p_k}, \quad m = 1, \ldots, K, \quad (20)
\]

and

\[
E \ln \sum_{k=1}^{K} \left[ \rho q_k(s) + (1 - \rho) R_k(s) \right] \frac{\lambda_k^*}{p_k} \geq 0. \quad (21)
\]

Condition (18) is necessary for this inequality to be strict (the “only if” part in (18)) because \( p_k = \lambda_k^* \) for all \( k = 1, \ldots, K \) implies that the left-hand side of (21) is zero.

2nd step. We fix the argument \( s \) and omit it in the notation. Consider the \( K \times K \) matrix

\[
A = (a_{mk}), \quad a_{mk} = \delta_{mk} - \rho \sum_{i=1}^{I} \lambda_m^i \lambda_k^i r^i
\]

where \( \delta_{mk} = 1 \) if \( m = k \) and \( \delta_{mk} = 0 \) if \( m \neq k \). Put

\[
b = (b_1, \ldots, b_K), \quad b_m = (1 - \rho) \sum_{k=1}^{K} R_k \frac{\sum_{i=1}^{I} \lambda_m^i \lambda_k^i r^i}{p_k}. \quad (22)
\]

Then (for the fixed \( s \)) the system of equations (20) can be written

\[
Aq = b \quad (23)
\]

\([q = q(s)]\). Indeed, the \( m \)th coordinate \((Aq - b)_m\) of the vector \( Aq - b \) can be expressed as follows

\[
(Aq - b)_m = \sum_{k=1}^{K} a_{mk} q_k - b_m = q_m - \rho \sum_{k=1}^{K} q_k \frac{\sum_{i=1}^{I} \lambda_m^i \lambda_k^i r^i}{p_k}.
\]
\[(1 - \rho) \sum_{k=1}^{K} R_k \frac{\sum_{i=1}^{I} \lambda_m^i \lambda_k^i r^i}{p_k} = q_m - \sum_{k=1}^{K} [\rho q_k + (1 - \rho) R_k] \frac{\sum_{i=1}^{I} \lambda_m^i \lambda_k^i r^i}{p_k},\]

which is equal to the difference between the left-hand side and the right-hand side of (20).

We can represent the matrix \( A \) as

\[ A = \text{Id} - \rho C, \]

where \( \text{Id} \) is the identity matrix and

\[ C = (c_{mk}), \quad c_{mk} = \frac{\sum_{i=1}^{I} \lambda_m^i \lambda_k^i r^i}{p_k}. \]

The norm of the linear operator \( C \) is not greater than one, because

\[ |Cx| = \sum_{m=1}^{K} \left| \sum_{k=1}^{K} c_{mk} x_k \right| \leq \sum_{m=1}^{K} \sum_{k=1}^{K} c_{mk} |x_k| = |x|, \quad (24) \]

where

\[ \sum_{m=1}^{K} c_{mk} = \sum_{m=1}^{K} \sum_{i=1}^{I} \lambda_m^i \lambda_k^i r^i = \frac{\sum_{i=1}^{I} \lambda_k^i r^i}{p_k} = 1. \quad (25) \]

Consequently, the operator \( \rho C \) is contracting, and so each of the equivalent equations

\[ Aq = 0, \quad q = \rho C q \]

has a unique solution. Thus the matrix \( A \) is non-degenerate, and the solution to the linear system (23) can be represented as

\[ q = A^{-1} b. \]

3rd step. Define

\[ c_k = \rho \frac{\lambda_k^*}{p_k}, \quad d = (1 - \rho) \sum_{k=1}^{K} R_k \frac{\lambda_k^*}{p_k}. \quad (26) \]

Then we have

\[ \langle c, q \rangle + d = \sum_{k=1}^{K} [\rho q_k + (1 - \rho) R_k] \frac{\lambda_k^*}{p_k}. \quad (27) \]
This expression appears in (21), and our goal is to estimate the expected logarithm of it (when $R_k$ and $q$ depend on $s$). To this end we write

$$\langle c, q \rangle = \langle c, A^{-1} b \rangle = \langle (A^{-1})' c, b \rangle = \langle (A')^{-1} c, b \rangle,$$  \hspace{1cm} (28)

where $A'$ denotes the conjugate matrix. In (28), we use the following identity

$$(A^{-1})' = (A')^{-1},$$

holding for each invertible linear operator $A$.

By virtue of (28),

$$\langle c, q \rangle = \langle b, l \rangle,$$  \hspace{1cm} (29)

where $l = (A')^{-1} c$, i.e., the vector $l$ is the solution to the linear system

$$A' l = c.$$

The matrix $A'$ is given by

$$A' = (d_{km}'), \quad d_{km}' = a_{mk} = \delta_{mk} - \rho \frac{\sum_{i=1}^{I} \lambda^*_m \lambda^*_k r^{*i}}{p_k},$$

and the linear system $A' l = c$ can be written

$$\sum_{m=1}^{K} (\delta_{mk} - \rho \frac{\sum_{i=1}^{I} \lambda^*_m \lambda^*_k r^{*i}}{p_k}) l_m = \rho \frac{\lambda^*_k}{p_k}, \quad k = 1, \ldots, K,$$

(see (26)) or equivalently,

$$l_k = \rho \left( \sum_{m=1}^{K} \frac{\sum_{i=1}^{I} \lambda^*_m \lambda^*_k r^{*i}}{p_k} l_m + \frac{\lambda^*_k}{p_k} \right), \quad k = 1, \ldots, K.$$  \hspace{1cm} (30)

Further, in view of (26) and (22), we obtain

$$d + \langle l, b \rangle = (1 - \rho) \sum_{k=1}^{K} R_k \frac{\lambda^*_k}{p_k} + (1 - \rho) \sum_{m=1}^{K} l_m \sum_{k=1}^{K} R_k \frac{\sum_{i=1}^{I} \lambda^*_m \lambda^*_k r^{*i}}{p_k} =$$

$$= (1 - \rho) \sum_{k=1}^{K} R_k \frac{\lambda^*_k}{p_k} + (1 - \rho) \sum_{k=1}^{K} R_k \sum_{m=1}^{K} l_m \frac{\sum_{i=1}^{I} \lambda^*_m \lambda^*_k r^{*i}}{p_k} =$$

18
\[(1 - \rho) \sum_{k=1}^{K} R_k \left[ \sum_{m=1}^{K} \frac{l_m \lambda_m^i \lambda_k^{*i}}{p_k} + \frac{\lambda_k^*}{p_k} \right] = \frac{(1 - \rho)}{\rho} \sum_{k=1}^{K} R_k l_k, \quad (31)\]

where the last equality follows from (30). Consequently,

\[\langle c, q \rangle + d = \langle l, b \rangle + d = \frac{(1 - \rho)}{\rho} \sum_{k=1}^{K} R_k l_k \quad (32)\]

(see (27) and (29)).

4th step. According to Step 1 of the proof, we have to establish inequality (21) for every solution \(q(s), \ s \in S\), of system (20) and show that this inequality is strict if

\[(p_1, \ldots, p_K) \neq (\lambda_1^*, \ldots, \lambda_K^*). \quad (33)\]

The considerations presented in Steps 2 and 3, allow to reduce this problem to the following one: for the solution \(l = (l_1, \ldots, l_K)\) to system (30), show that

\[E \ln \left( \frac{(1 - \rho)}{\rho} \sum_{k=1}^{K} R_k(s) l_k \right) \geq 0 \]

(see (32) and (29)). Additionally, it has to be shown that the last inequality is strict if assumption (33) holds. The advantage of the new problem comparative to the original one lies in the fact that system (30), in contrast with (20), does not depend on \(s\).

We write (30) equivalently as

\[\frac{(1 - \rho)}{\rho} p_k l_k = \rho \sum_{m=1}^{K} p_m \frac{(1 - \rho)}{\rho} l_m \sum_{i=1}^{l} \frac{\lambda_m^i \lambda_k^{*i}}{p_m} + \frac{(1 - \rho)}{\rho} \lambda_k^*, \]

and, by changing variables

\[f_k = \frac{(1 - \rho)}{\rho} l_k p_k, \]

we transform (30) to

\[f_k = \rho \sum_{m=1}^{K} f_m \frac{\sum_{i=1}^{l} \lambda_m^i \lambda_k^{*i}}{p_m} + (1 - \rho) \lambda_k^*, \quad k = 1, \ldots, K. \quad (34)\]

Then

\[\frac{(1 - \rho)}{\rho} \sum_{k=1}^{K} R_k l_k = \sum_{k=1}^{K} \frac{R_k f_k}{p_k}, \]

19
and the problem reduces to the following one: given the solution \((f_1, ..., f_K)\) to system (34), show that

\[
E \ln \sum_{k=1}^{K} R_k(s) \frac{f_k}{p_k} \geq 0
\]  

(35)

with strict inequality if \((p_1, ..., p_K) \neq (\lambda_1^*, ..., \lambda_K^*)\).

Note that the affine operator defined by the right-hand side of (34) is contracting (see (24) and (25)) and leaves the non-negative cone \(\mathbb{R}_+^K\) invariant. Therefore there exists a unique vector \(f = (f_1, ..., f_K)\) solving (34). Furthermore, this vector is strictly positive (which follows from the strict positivity of \(\lambda^* = (\lambda_1^*, ..., \lambda_K^*)\)) and satisfies \(\sum_{k=1}^{K} f_k = 1\). The last equality can be obtained by summing up equations (34) over \(k = 1, ..., K\).

By virtue of Jensen’s inequality, applied to the concave function \(\ln(\cdot)\), we have

\[
E \ln \sum_{k=1}^{K} R_k(s) \frac{f_k}{p_k} \geq E \sum_{k=1}^{K} R_k(s) \ln \frac{f_k}{p_k} = \sum_{k=1}^{K} \lambda_k^* \ln \frac{f_k}{p_k}.
\]

(We use here the fact that \(\sum_{k=1}^{K} R_k(s) = 1\) for all \(s\).) Thus it is sufficient to prove that if a vector \((f_1, ..., f_K)\) satisfies (34), then

\[
\sum_{k=1}^{K} \lambda_k^* \ln \frac{f_k}{p_k} \geq 0,
\]

(36)

and inequality (36) is strict when assumption (33) is fulfilled. This problem is purely deterministic: no random parameter \(s\) is involved either in (34) or in (36).

5th step. Put \(g_k = \frac{f_k}{p_k}, \ k = 1, ..., K\). Then, from (34), we get

\[
p_kg_k = \rho \sum_{m=1}^{K} g_m \sum_{i=1}^{l} \lambda_m^i \lambda_k^i r^i + (1 - \rho) \lambda_k^*, \ k = 1, ..., K.
\]

(37)

Let us multiply both sides of these equations by \(\ln g_k\) and sum up over \(k = 1, ..., K\):

\[
\sum_{k=1}^{K} p_k g_k \ln g_k = \rho \sum_{k=1}^{K} (\ln g_k) \sum_{m=1}^{K} g_m \sum_{i=1}^{l} \lambda_m^i \lambda_k^i r^i + (1 - \rho) \sum_{k=1}^{K} \lambda_k^* \ln g_k.
\]
This yields
\[
\sum_{k=1}^{K} \lambda_k^* \ln g_k = \frac{1}{1 - \rho} \left[ \sum_{k=1}^{K} p_k g_k \ln g_k - \rho \sum_{k=1}^{K} (\ln g_k) \sum_{m=1}^{K} g_m \sum_{i=1}^{l} \lambda_m^i \lambda_k^i r^i \right].
\]

Further, we have
\[
\sum_{k=1}^{K} \lambda_k^* \frac{f_k}{p_k} = \sum_{k=1}^{K} \lambda_k^* \ln g_k
\]
(recall that \( g_k = f_k/p_k \)). Thus, in order to prove the desired inequality (36) it is sufficient to verify the relation
\[
\sum_{k=1}^{K} p_k g_k \ln g_k - \rho \sum_{k=1}^{K} (\ln g_k) \sum_{m=1}^{K} g_m \sum_{i=1}^{l} \lambda_m^i \lambda_k^i r^i \geq 0. \tag{38}
\]

If inequality (38) is strict, then (36) is strict as well.

We have
\[
\sum_{k=1}^{K} p_k g_k \ln g_k = \sum_{k=1}^{K} f_k \ln \frac{f_k}{p_k} \geq 0, \tag{39}
\]
by virtue of the well-known inequality (recall that \( f, p \in \Delta^K \))
\[
\sum_{k=1}^{K} f_k \ln f_k \geq \sum_{k=1}^{K} f_k \ln p_k, \tag{40}
\]
which is strict if
\[
(f_1, \ldots, f_K) \neq (p_1, \ldots, p_K). \tag{41}
\]
Therefore relation (38) is valid if
\[
\sum_{k=1}^{K} (\ln g_k) \sum_{m=1}^{K} g_m \sum_{i=1}^{l} \lambda_m^i \lambda_k^i r^i \leq 0. \tag{42}
\]

In the rest of the proof, we will assume that the opposite inequality holds:
\[
\sum_{k=1}^{K} (\ln g_k) \sum_{m=1}^{K} g_m \sum_{i=1}^{l} \lambda_m^i \lambda_k^i r^i > 0. \tag{43}
\]
Then (38) will be obtained if we establish that

\[ \sum_{k=1}^{K} p_k g_k \ln g_k \geq \sum_{k=1}^{K} (\ln g_k) \sum_{m-1}^{K} g_m \sum_{i-1}^{I} \lambda_m^i \lambda_k^i r^i. \] (44)

Indeed, then we have

\[ \sum_{k=1}^{K} p_k g_k \ln g_k \geq \sum_{k=1}^{K} (\ln g_k) \sum_{m-1}^{K} g_m \sum_{i-1}^{I} \lambda_m^i \lambda_k^i r^i \geq \rho \sum_{k=1}^{K} (\ln g_k) \sum_{m-1}^{K} g_m \sum_{i-1}^{I} \lambda_m^i \lambda_k^i r^i, \]

which yields (38). In the above chain of relations, the last equality holds by virtue of (43).

To verify inequality (44) we write

\[ \sum_{k=1}^{K} p_k g_k \ln g_k = \sum_{k=1}^{K} \sum_{i-1}^{I} \lambda_k^i r^i g_k \ln g_k = \sum_{i-1}^{I} r^i \sum_{k=1}^{K} \lambda_k^i g_k \ln g_k, \]

and

\[ \sum_{k=1}^{K} (\ln g_k) \sum_{m-1}^{K} g_m \sum_{i-1}^{I} \lambda_m^i \lambda_k^i r^i = \sum_{i-1}^{I} r^i \sum_{k=1}^{K} \sum_{m-1}^{K} (\lambda_k^i \ln g_k)(g_m \lambda_m^i). \]

Thus to prove (44) it remains to check that

\[ \sum_{k=1}^{K} \lambda_k^i g_k \ln g_k \geq (\sum_{k=1}^{K} \lambda_k^i \ln g_k)(\sum_{k=1}^{K} g_k \lambda_k^i) \] (45)

for each \( i = 1, \ldots, I. \)

Let us fix \( i \) and put \( \lambda_k = \lambda_k^i. \) Inequality (45) can be written

\[ \mathbb{E}[g \ln g] \geq (\mathbb{E} \ln g)\mathbb{E}g, \]

where “\( \mathbb{E} \)” stands for the weighted average

\[ \mathbb{E}g = \sum_{k=1}^{K} g_k \lambda_k \quad [\lambda_k > 0, \sum_{k=1}^{K} \lambda_k = 1]. \]

Observe that the function \( \phi(g) = g \ln g \) is strictly convex. Consequently,

\[ \mathbb{E}\phi(g) \geq \phi(\mathbb{E}g), \] (46)
and the inequality is strict if \( g_k \neq g_m \) for some \( k \) and \( m \). Thus

\[
\mathbb{E}[g \ln g] \geq (\mathbb{E}g) \ln \mathbb{E}g \geq (\mathbb{E}g)(\mathbb{E}\ln g),
\]

(47)

where the former inequality in this chain of relations coincides with (46) and the latter is a consequence of the concavity of the function \( \ln(\cdot) \). Furthermore, both inequalities in (47) are strict provided that \( g_k \neq g_m \) for some \( k \). If the last condition does not hold, then

\[
\frac{f_k}{p_k} = c
\]

for some constant \( c \), which must necessarily be equal to one because \( \sum f_k = \sum p_k = 1 \). Thus if \( g_k = g_m \) for all \( k, m \), then \( f_k = p_k, k = 1, 2, ..., K \), which implies (see below) that \( p_k = \lambda^*_k \) for all \( k \).

6th step. At the previous step of the proof, we established inequality (38) and hence (36). Moreover, the arguments conducted show that inequality (38) (and hence (36)) is strict if condition (41) is fulfilled. Indeed, if relation (42) holds then, under assumption (41), we have a strict inequality in (39), which implies a strict inequality in (38). Alternatively, if relation (43), opposite to (42), holds, then strict inequalities in (47) and (45) imply strict inequalities in (44) and (38).

Thus to complete the proof it suffices to show that if

\[
f_k = p_k, \ k = 1, 2, ..., K,
\]

then

\[
p_k = \lambda^*_k, \ k = 1, 2, ..., K.
\]

Indeed, if \( f_k = p_k \), then we have

\[
p_k = \rho \sum_{m=1}^{K} p_m \frac{\sum_{i=1}^{I} \lambda_m^i \gamma^I x^i}{p_m} + (1 - \rho) \lambda^*_k, \ k = 1, ..., K,
\]

which implies

\[
p_k = \rho \sum_{m=1}^{K} \sum_{i=1}^{I} \lambda_m^i \gamma^I x^i + (1 - \rho) \lambda^*_k = \rho p_k + (1 - \rho) \lambda^*_k, \ k = 1, ..., K.
\]

Thus \( (1 - \rho)p_k = (1 - \rho)\lambda^*_k \), and so \( p_k = \lambda^*_k \). \( \square \)
6 The Kelly portfolio and the market portfolio

According to Theorem 2, the expected logarithmic return on the Kelly portfolio \((\lambda_1^*, ..., \lambda_K^*)\) is non-negative. It is strictly positive if and only if the market portfolio \((p_1, ..., p_K)\) does not coincide with \((\lambda_1^*, ..., \lambda_K^*)\). Of course it can happen at some moment of time that \((\lambda_1^*, ..., \lambda_K^*) = (p_1, ..., p_K)\). But can it happen that the market portfolio coincides with \((\lambda_1^*, ..., \lambda_K^*)\) at two consecutive moments of time? In other words, can the system of equalities

\[
q_k(s) = p_k = \lambda_k^* \quad (k = 1, ..., K, \ s \in S)
\] (48)

hold? Recall that we denote by \(p_k\) the price of the asset \(k\) corresponding to the vector \(r = (r^1, ..., r^I)\) of relative wealth at some fixed moment of time,

\[
p_k = \langle \lambda_k, r \rangle = \sum_{i=1}^{I} \lambda_k^i r^i
\]

and by \(q_k(s)\) the price of the asset at the next moment of time, when the state of the world realized is \(s\):

\[
q_k(s) = \langle \lambda_k, F(s, r) \rangle = \sum_{i=1}^{I} \lambda_k^i F^i(s, r).
\]

The question we formulated is important for the analysis of the asymptotic behavior of the relative wealth of an investor using the Kelly rule. As Proposition 4 below shows, the answer to this question (under the assumptions we impose) is negative.

Recall that we assume that there are no redundant assets, i.e., the functions \(R_1(s), ..., R_K(s)\) are linearly independent. This assumption will be used in the following proposition.

**Proposition 4** Suppose one of the following assumptions is fulfilled.

(a) All the investors \(i = 1, 2, ..., I\) use portfolio rules \(\lambda^i = (\lambda_1^i, ..., \lambda_K^i)\) distinct from the Kelly rule \(\lambda^* = (\lambda_1^*, ..., \lambda_K^*)\).

(b) All the investors \(i = 2, 3, ..., I\) use portfolio rules \(\lambda^i = (\lambda_1^i, ..., \lambda_K^i)\) distinct from the Kelly rule \(\lambda^* = (\lambda_1^*, ..., \lambda_K^*)\), and the wealth share \(r^1\) of investor 1 is less than one.

Then equations (48) cannot hold.
Proof. The variables $q_k(s)$, $p_k$ and $r_k$ ($k = 1, ..., K$) are related to each other by the system of equations (20). Suppose equations (48) hold. Then, from (20), we obtain:

$$\lambda^*_m = \sum_{k=1}^{K} [\rho \lambda^*_k + (1 - \rho) R_k(s)] \frac{\sum_{i=1}^{l} \lambda^*_m \lambda^*_k r^i}{\lambda^*_k}, \quad m = 1, ..., K,$$

or equivalently,

$$\lambda^*_m = \sum_{k=1}^{K} \bar{R}_k(s) \sum_{i=1}^{l} \frac{\lambda^*_m \lambda^*_k r^i}{\lambda^*_k}, \quad m = 1, ..., K, \quad (49)$$

where

$$\lambda^*_k = E\bar{R}_k(s), \quad \bar{R}_k(s) = \rho \lambda^*_k + (1 - \rho) R_k(s).$$

Observe that if there are no redundant assets, then the relation

$$\sum_{k=1}^{K} \gamma_k \bar{R}_k(s) = 0$$

implies $\gamma_1 = ... = \gamma_K = 0$. Indeed, suppose that

$$\sum_{k=1}^{K} \gamma_k [\rho \lambda^*_k + (1 - \rho) R_k(s)] = 0. \quad (50)$$

Then we have

$$0 = E \sum_{k=1}^{K} \gamma_k [\rho \lambda^*_k + (1 - \rho) R_k(s)] = \sum_{k=1}^{K} \gamma_k [\rho \lambda^*_k + (1 - \rho) \lambda^*_k] = \sum_{k=1}^{K} \gamma_k \lambda^*_k,$$

which in view of (50) yields

$$\sum_{k=1}^{K} \gamma_k R_k(s) = - \frac{\rho}{1 - \rho} \sum_{k=1}^{K} \gamma_k \lambda^*_k = 0,$$

and so $\gamma_1 = ... = \gamma_K = 0$ because the functions $R_k(\cdot)$, $k = 1, ..., K$, are linearly independent.

From formula (49) and the relation

$$\lambda^*_m = p_m = \sum_{i=1}^{l} \lambda^*_m r^i,$$

we get...
we obtain
\[ \sum_{i=1}^{I} \lambda_m^i r^i = \sum_{k=1}^{K} \tilde{R}_k(s) \sum_{i=1}^{I} \frac{\lambda_m^i \lambda_k^i r^i}{\lambda_k^i}, \quad m = 1, \ldots, K. \] (51)

We have \( \sum_{k=1}^{K} \tilde{R}_k(s) = 1 \), and so equations (51) imply
\[ \sum_{k=1}^{K} \tilde{R}_k(s) \gamma_k^m = 0, \quad m = 1, \ldots, K, \]
where
\[ \gamma_k^m = \frac{1}{\lambda_k^*} \sum_{i=1}^{I} \lambda_m^i \lambda_k^* r^i - \sum_{i=1}^{I} \lambda_m^i r^i. \]

Since there are no redundant assets, we have \( \gamma_k^m = 0 \) for each \( m \) and \( k \). This gives
\[ \sum_{i=1}^{I} \lambda_m^i \lambda_k^i r^i - \lambda_k^* \sum_{i=1}^{I} \lambda_m^i r^i = 0, \quad k, m = 1, \ldots, K, \]
which can be written as
\[ \sum_{i=1}^{I} \lambda_m^i (\lambda_k^i - \lambda_k^*) r^i = 0, \quad k, m = 1, \ldots, K. \]

We derive two expressions from this equation. The first by setting \( k = m \) in the foregoing formula. The second by adding up over \( m = 1, \ldots, K \). We find
\[ \sum_{i=1}^{I} \lambda_k^i (\lambda_k^i - \lambda_k^*) r^i = 0, \quad k = 1, \ldots, K, \quad \text{and} \quad \sum_{i=1}^{I} (\lambda_k^i - \lambda_k^*) r^i = 0. \]

Multiplying the second equation by \(-\lambda_k^*\) and adding it up with the first, we obtain
\[ 0 = \sum_{i=1}^{I} [\lambda_k^i (\lambda_k^i - \lambda_k^*) r^i - \lambda_k^* (\lambda_k^i - \lambda_k^*) r^i] = \sum_{i=1}^{I} (\lambda_k^i - \lambda_k^*)^2 r^i. \]

Consequently, we have
\[ (\lambda_k^i - \lambda_k^*)^2 r^i = 0, \quad i = 1, \ldots, I, \quad k = 1, \ldots, K. \] (52)
Suppose condition (a) holds. Since \( \sum_{i=1}^{I} r^i = 1 \) and \( r^i \geq 0 \), we have \( r^j > 0 \) for some \( j = 1, \ldots, I \). Then, from (52), we get
\[
\lambda_k^j - \lambda_k^* = 0, \ k = 1, \ldots, K,
\]
which is a contradiction.

If condition (b) is fulfilled, then \( \sum_{i=2}^{I} r^i = 1 - r^1 \), and so \( r^j > 0 \) for some \( j = 2, \ldots, I \). This implies (53), and the contradiction obtained completes the proof. \( \square \)

7 Limiting behavior of the Kelly investor’s relative wealth

Let \( r_0 \) be a strictly positive vector in \( \Delta^I \). Define recursively the sequence of random vectors \( r_0, r_1(s^1), r_2(s^2), \ldots \) by the formula \( r_t = F(s_t, r_{t-1}) \). Then \( r_t = (r_t^1, \ldots, r_t^I) \) is the vector of relative wealths of the investors \( i = 1, \ldots, I \) at date \( t \), depending on the realization \( s^t = (s_1, \ldots, s_t) \) of states of the world. It follows from Proposition 3 that \( r_t > 0 \) as long as \( r_{t-1} > 0 \) and so all the vectors \( r_t(s^t) \) are strictly positive for all \( t \) and \( s^t \). Consequently, the random variables
\[
\ln r_t^i = \ln r_t^i(s^t), \ i = 1, \ldots, I, \ t = 0, 1, \ldots
\]
are well-defined and finite. Clearly, they have finite expectations because each of them takes on a finite number of values (since the set \( S \) is finite).

Suppose investor 1 uses the Kelly rule
\[
\lambda^* = (\lambda_1^*, \ldots, \lambda_K^*) = (ER_1(s), \ldots, ER_K(s)).
\]
Consider the growth rate \( r_{t+1}^1/r_t^1 \) of investor 1’s relative wealth. It can be expressed as follows:
\[
\frac{r_{t+1}^1}{r_t^1} = g^1(s_{t+1}, r_t) = \frac{F^1(s_{t+1}, r_t)}{r_t^1} \quad [r_t = r_t(s^t)]
\]
(see (13)), and since the strategy \( \lambda^1 \) of investor 1 coincides with the Kelly rule \( \lambda^* \), we have
\[
\frac{r_{t+1}^1}{r_t^1} = g^1(s_{t+1}, r_t) = f(s_{t+1}, r_t),
\]
where \( f(s, r) \) is the function defined by (15).
Denote by $\xi_t = \xi_t(s^t)$ the logarithm of the relative wealth of investor 1,

$$\xi_t = \ln r_t^1.$$  

We claim that the sequence $\xi_t$ is a submartingale:

$$E(\xi_{t+1}|s^t) \geq \xi_t. \quad (55)$$  

Indeed, we have

$$E(\xi_{t+1}|s^t) - \xi_t = E[(\xi_{t+1} - \xi_t)|s^t] = E[(\ln r_{t+1}^1 - \ln r_t^1)|s^t] =$$

$$E[(\ln \frac{r_{t+1}^1}{r_t^1})|s^t] = E[\ln f(s_{t+1}, r_t)|s^t] = E[\ln f(s, r_t)|r_t = r_t(s^t) =$$

$$\sum_{a \in S} \pi(s) \ln f(s, r_t(s^t)), $$

where $\pi(s) > 0$ is the probability that $s_{t+1} = s$. The last two equalities in the above chain of relations follow from the fact that the random variables $s_1, s_2, \ldots$ are independent and identically distributed. By virtue of Theorem 2,

$$\sum_{a \in S} \pi(s) \ln f(s, r_t(s^t)) \geq 0, $$

which proves (55). Since $0 < r_t^1 \leq 1$, we have $\xi_t \leq 0$, and so $\xi_t, t = 0, 1, \ldots,$ is a non-positive submartingale. As is well-known, a non-positive submartingale converges almost surely (a.s.)

$$\xi_t \to \xi_\infty \text{ (a.s.) as } t \to \infty$$

(see, e.g., [21], Section IV.5). This implies

$$r_t^1 = e^{\xi_t} \to e^{\xi_\infty} > 0 \text{ (a.s.).}$$

This leads to the following result.

**Theorem 3** The relative wealth of a Kelly investor converges a.s., and the limit is strictly positive.

It follows from Theorem 3 that an investor using the Kelly strategy survives with probability one. A key result of this study is Theorem 4 below, asserting that if one of the investors uses the Kelly rule and all the others use other strategies, distinct from the Kelly one, then the Kelly investor is the only survivor in the market selection process.
Theorem 4 Let the strategy of investor 1 coincide with the Kelly rule: \( \lambda_k^1 = \lambda_k^*, \ k = 1, \ldots, K \). Let the strategies of investors \( i = 2, \ldots, I \) be distinct from the Kelly rule:

\[
(\lambda_i^1, \ldots, \lambda_i^K) \neq (\lambda_1^*, \ldots, \lambda_K^*).
\]

Then the relative wealth \( r_t^1 \) of investor 1 converges to one almost surely.

We note that if \( \rho = 0 \), Theorem 4 follows from the main result of the paper [12]. Methods developed in this work are different in some respects from those in [12].

Proof of Theorem 4. By virtue of Theorem 3, the limit \( r_\infty^1 := \lim r_t^1 \) exists a.s. and is strictly positive. Suppose the assertion we wish to prove is not valid. Then we have

\[
P\{0 < \lim r_t^1 < 1\} > 0. \tag{56}
\]

Let us write for shortness \( E_t(\cdot) \) in place of \( E(\cdot|s^t) \). If \( \xi_t \) is a non-positive submartingale, then \( E_{t-1}\xi_{t+1} - \xi_{t-1} \to 0 \) a.s. (see Lemma 1 in the Appendix). By applying this fact to the non-positive submartingale \( \xi_t = \ln r_t^1 \), we obtain

\[
E_{t-1}(\ln \frac{r_t^1}{r_{t-1}^1} + \ln \frac{r_{t+1}^1}{r_t^1}) = E_{t-1}\ln \frac{r_{t+1}^1}{r_{t-1}^1} = E_{t-1}\xi_{t+1} - \xi_{t-1} \to 0 \text{ (a.s.)}. \tag{57}
\]

By using the fact that the random elements \( s^{t-1}, s_t \) and \( s_{t+1} \) are independent and representing the histories \( s^t, s^{t+1} \) as

\[
s^t = (s^{t-1}, s_t), \ s^{t+1} = (s^{t-1}, s_t, s_{t+1}),
\]

we get

\[
E_{t-1}(\ln \frac{r_t^1}{r_{t-1}^1} + \ln \frac{r_{t+1}^1}{r_t^1}) = E[\ln \frac{r_t^1}{r_{t-1}^1}|s^{t-1}] + E[\ln \frac{r_{t+1}^1}{r_t^1}|s^{t-1}] =
\]

\[
\sum_{s \in S} P\{s_t = s\} \ln \frac{r_t^1(s^{t-1}, s)}{r_{t-1}^1(s^{t-1})} + \sum_{s \in S} P\{s_t = s\} \sum_{\sigma \in S} P\{s_{t+1} = \sigma\} \ln \frac{r_{t+1}^1(s^{t-1}, s, \sigma)}{r_t^1(s^{t-1}, s)} =
\]

\[
\sum_{s \in S} \pi(s) \ln \frac{r_t^1(s^{t-1}, s)}{r_{t-1}^1(s^{t-1})} + \sum_{s \in S} \pi(s) \sum_{\sigma \in S} \pi(\sigma) \ln \frac{r_{t+1}^1(s^{t-1}, s, \sigma)}{r_t^1(s^{t-1}, s)} =
\]

\[
\sum_{s \in S} \pi(s) \ln f(s, r_{t-1}(s^{t-1})) + \sum_{s \in S} \pi(s) \sum_{\sigma \in S} \pi(\sigma) \ln f(\sigma, r_t(s^{t-1}, s)) =
\]

29
\[
\sum_{s \in S} \pi(s) \ln f(s, r_{t-1}(s^{t-1})) + \sum_{s \in S} \pi(s) \sum_{\sigma \in S} \pi(\sigma) \ln f(\sigma, F(s, r_{t-1}(s^{t-1}))) \quad \text{for all } t \geq 1.
\] 

The last but one equality in the above chain of relations is valid because the strategy of investor 1 coincides with the Kelly rule \((\lambda^1 = \lambda^*)\) and the last equality holds because \(r_t(s^{t-1}, s) = F(s, r_{t-1}(s^{t-1}))\).

By virtue of (56), (57) and (58), there exists a realization \((s_1, ..., s_t, ...)\) of the process of states of the world such that, for the sequence of vectors \(r_{t-1} = r_{t-1}(s^{t-1}) \in \Delta^I\), we have

\[
0 < \lim_{t \to \infty} r_{t-1}^1 < 1, \tag{59}
\]

\[
\sum_{s \in S} \pi(s) \ln f(s, r_{t-1}) + \sum_{s \in S} \pi(s) \sum_{\sigma \in S} \pi(\sigma) \ln f(\sigma, F(s, r_{t-1})) \to 0. \tag{60}
\]

In the rest of the proof, we will fix such a realization \((s_1, ..., s_t, ...)\) and write \(r_{t-1}\) in place of \(r_{t-1}(s^{t-1})\).

Since the simplex \(\Delta^I\) is compact, there exists a sequence \(t_1 < t_2 < ...\) and a vector \(r \in \Delta^I\) such that

\[
r_{t_n-1} \to r \in \Delta^I. \tag{61}
\]

It follows from (59) and (61) that the first coordinate \(r^1\) if the vector \(r = (r^1, ..., r^I)\) satisfies

\[
0 < r^1 < 1. \tag{62}
\]

Relations (60) and (61) imply

\[
\sum_{s \in S} \pi(s) \ln f(s, r) + \sum_{s \in S} \pi(s) \sum_{\sigma \in S} \pi(\sigma) \ln f(\sigma, F(s, r)) = 0 \tag{63}
\]

because the function \(\ln f(s, r) = \ln g^1(s, r)\) is continuous in \(r \in \Delta^I\) (see Proposition 3).

By virtue of Theorem 2,

\[
\sum_{s \in S} \pi(s) \ln f(s, r) \geq 0, \quad \sum_{\sigma \in S} \pi(\sigma) \ln f(\sigma, F(s, r)) \geq 0 \quad \text{for all } s \in S.
\]

Consequently, it follows from (63) that

\[
\sum_{s \in S} \pi(s) \ln f(s, r) = 0, \tag{64}
\]

\[30\]
and

$$
\sum_{\sigma \in S} \pi(\sigma) \ln f(\sigma, F(s, r)) = 0 \text{ for all } s \in S. \quad (65)
$$

According to Theorem 2, relation (64) can hold only if

$$
\langle \lambda_k, r \rangle = \lambda_k^*, \quad k = 1, \ldots, K, \quad (66)
$$

and equations (65) imply

$$
\langle \lambda_k, F(s, r) \rangle = \lambda_k^*, \quad k = 1, \ldots, K, \quad s \in S. \quad (67)
$$

By virtue of Proposition 4, relations (62), (66) and (67) cannot hold simultaneously. This is a contradiction.

\[ \square \]

A Appendix: A lemma about submartingales

**Lemma 1** Let $\xi_t$ be a non-positive submartingale. Then the sequence of non-negative random variables $\zeta_t = E_{t-1} \xi_{t+1} - \xi_{t-1}$ converges to zero a.s.

**Proof.** We have $\zeta_t \geq 0$ by the definition of a submartingale. Further, $E\xi_t = (E\xi_{t+1} - E\xi_t) + (E\xi_t - E\xi_{t-1})$, and so

$$
\sum_{t=1}^{N} E\xi_t = \sum_{t=1}^{N} (E\xi_{t+1} - E\xi_t) + \sum_{t=1}^{N} (E\xi_t - E\xi_{t-1}) =
$$

$$
E\xi_{N+1} - E\xi_1 + E\xi_N - E\xi_0 \leq -E\xi_1 - E\xi_0
$$

because $E\xi_t \leq 0$ for each $t$. Therefore the series of the expectations $\sum_{t=1}^{\infty} E\xi_t$ of the non-negative random variables $\zeta_t$ converges, which implies (see, e.g., Corollary to Theorem 11, in Chapter VI in [20]) that $\zeta_t \to 0$ (a.s.).

\[ \square \]

References


