International Journal of Theoretical and Applied Finance
Vol. 4, No. 2 (2001) 335–354

TRANSACTION COSTS: A NEW POINT OF VIEW

ROBERTO BAVIERA*
Departement de Finance & Economie, Groupe HEC,
F-78351 Jouy-en-Josas Cedex, France

Received 22 July 1999
Accepted 15 February 2000

We consider a new point of view to portfolio selection in presence of transaction costs which allows us to map the problem into one without costs. The proposed approach connects all the quantities of interest to the average lag between two trades and the probabilities of reaching a price fluctuation of a given size. This active portfolio managing in presence of costs is based on quantities directly measurable in real markets and it leads to analytic results in the Wiener case.

1. Introduction

Allocation of wealth among portfolios in the presence of costs is the everyday problem for a trader (or a generic investor) in a financial market. The two relevant questions in active portfolio managing are

- when should the trader rebalance his portfolio?
- how many assets should he sell or buy?

The aim of this paper is to address these questions in the elementary case of geometric Brownian motion prices for a trader with an infinite horizon. However the proposed methodology is far more general: we discuss strength and limits of the approach in the case of prices following a more generic stationary stochastic process.

A trader knows that in a perfect financial world the optimal strategy is to modify his portfolio accordingly to price changes. After every small price modification his asset allocation is not optimal anymore and he should sell or buy a small number of the assets in his portfolio to reproduce the optimal allocation. This has led to the well-known continuous time theory of finance that can be found in excellent monographs [1], [2] and [3].

However in a real world transaction costs exist, for the trader it is not worth continously changing his position but it is better to act only when prices change by a "relevant" amount. It is straightforward that the advantage of a portfolio

*E-mail: bavierar@hec.fr

modification should be at least equal to the cost paid. For example, an investor does not sell an asset if the bid-ask spread is larger than the difference between the selling price and the one he has paid to buy it.[a] In this case the benefit arising from the portfolio modification is surely lower than the costs due to the spread.

We consider this spread in asset price the only source of trading costs. This appears to be reasonable both for a portfolio manager who trades directly with a market maker or for a small investor who has to commit his trade to a broker. The spread includes not only the bid/ask difference of the market but also the commission fees (generally proportional to the amount of money involved in the trade) or other proportional costs paid by the investor.

In this paper we deal with an investment on a risk-less asset paying a fixed interest rate (e.g. bank account or cash invested in short term bonds) and on a risky asset (e.g. asset shares or a portfolio of financial assets).

This problem has already been tackled in the financial literature [4–7] for geometric Brownian motion prices. In particular the same kind of investment (risk-less and one risky asset) is addressed by Dumas et Luciano [6], who considers a trader with an utility in the HARA class[b] and by Taksar *et al.* [4], where the trader maximizes the almost sure long-run growth rate of the capital. The idea of both papers is that, if the price of the risky asset is described by a geometric Brownian motion, it is possible to write a continuous time Bellman–Hamilton–Jacobi equation for the function of capital to maximize, utility in one case, long-run growth rate in the other. In this equation there is a term related to *the* price at time t and a dissipation term which takes into account the costs paid up to that moment. It is then possible to solve (at least numerically) the partial differential equation via a *singular* control Brownian motion technique (see e.g. [9] and [10] for a review).

This approach is open to two main criticisms. The first obvious one is that if the price is not a geometric Brownian motion, nothing can be said about a "reasonable" (even if not optimal) trading strategy: it is possible to write the Bellman–Hamilton–Jacoby equation only in the geometric Brownian motion case. The other criticism is the way the costs are introduced in the model. Since one is considering a function related to an *absolute* value of the risky-asset, the trader has always to remember the capital spent up to that time in his trading strategy. In this way a non trivial memory is introduced in the problem, memory which was not present in the original model with independent returns.

We propose a new point of view to the optimal portfolio selection in presence of transaction costs, related to the way a trader behaves in practice.

First, we focus our attention only on the times investor changes his position and we ignore what happens between two trades. Every time he modifies his portfolio he has to know which fraction of his capital has to be invested in the risky asset,

[a] This is not any more true in the case the trader desires to close his position because of liquidity necessities or a private information of a subsequent fall of the asset price. Considering a stationary process for the price and an investor with infinite horizon such problems do not take place.
[b] See e.g. [8] for a list of the general characteristics of these utilities.

and when he should modify his position again: this is all the information needed for every practical purpose. We do not observe continously the process and in this way we avoid writing a partial differential equation: this allows us to generalize the approach to a large class of portfolio allocations.

Second, instead of adding costs to an abstract *absolute* value we consider an asset value *relative* to investor's behavior, i.e. the *true* price paid or received by the trader depending on the operation (ask or bid) he performs. In presence of costs the value is a matter of conventions: there are no financial reasons to prefer one particular choice. Of course the optimal strategy cannot depend on the convention used. We show that this convention maps the portfolio optimization problem in presence of transaction costs into a similar one without costs.

This new point of view allows us to face portfolio allocation in presence of costs in a quite general (and intuitive) way and to obtain analytical solutions for all the quantities of interest in the case of a geometric Brownian motion price.

The paper is organized as follows: in Sec. 2 we state the portfolio selection problem and describe the method. We summarize the case of absence of costs in a version appropriate for our purposes in Sec. 3. In Sec. 4 we consider the *relative* value approach to the problem in presence of costs: both the exact solution and an approximation of it are discussed in detail. Finally Sec. 5 is devoted to summarizing the results. In the appendix we deduce for a bounded Wiener process the probabilities to reach the boundaries and the average time needed.

2. The Method

We consider an investor with an infinite time horizon, who diversifies his wealth W in a portfolio with a risk-less bank account and a risky asset, e.g. stocks. He tries to implement a strategy to extract the whole information due to the knowledge of the stochastic process for price evolution.

A trading strategy is a sequence of trades[c], where a rule specifies the capital fraction hold in shares after a trade and it determines the next time the portfolio should be rebalanced. Before a rebalancing, the *value* of the risky-asset has changed by a given amount.

Asset *value* is well defined only when the asset is traded: in presence of transaction costs it is the bid price S^b when the asset is sold and the ask price S^a when it is bought. At time t bid and ask processes for the logarithm of the prices (hereafter log-prices) are defined respectively:

$$r_t^b \equiv \ln S_t^b \,, \quad r_t^a \equiv \ln S_t^a = r_t^b + \gamma \,,$$

where γ is the transaction cost. In this paper we consider the case of a time-independent γ. We call bid (ask) return the difference between bid (ask) log-prices at two different times.

[c]A dynamic portfolio managing is *always* a discrete time problem, continuous time results can be obtained as limit case whenever needed.

We assume the same *value* not only for the assets traded but also for the ones still in the portfolio of the investor. This convention of an asset value *relative* to trader's behavior seems the most natural (since all stocks of the same asset are equal) and eliminates the main difference between portfolio managing with and without costs: the presence of two prices for the risky asset.

Let us consider the kth modification of the investor's portfolio. At the kth trade a fraction l_k of the wealth W_k is invested in stocks and the remaining in the risk-free asset. Between the kth and the sequent trade the logarithmic variation of the *value* reaches

$$\begin{array}{ll} \text{either} & \Delta_k^+ > 0 \\ \text{or} & -\Delta_k^- < 0 \end{array}$$

(hereafter called barriers) and the lag between the two portfolio changes (usually called exit time in the mathematical literature [11]) is

$$\mathcal{T}_k \equiv t_{k+1} - t_k ,$$

where t_k is the time when the kth trade is realized. The *value* of the capital fraction invested in risky-assets is $\exp(z_k \Delta_k^{z_k}) l_k W_k$ just before the $(k+1)$th trade and $l_{k+1} W_{k+1}$ immediately after, where the random variable z_k is the sign of the return between times t_k and t_{k+1}.

As shown in Fig. 1 if an ask at kth trade is followed by a bid, an ask return of $\Delta_k^+ + \gamma$ produces a logarithmic variation of only Δ_k^+ in the *value* of the risky asset, the remaining γ is simply due to the fact that the trader cannot sell at the

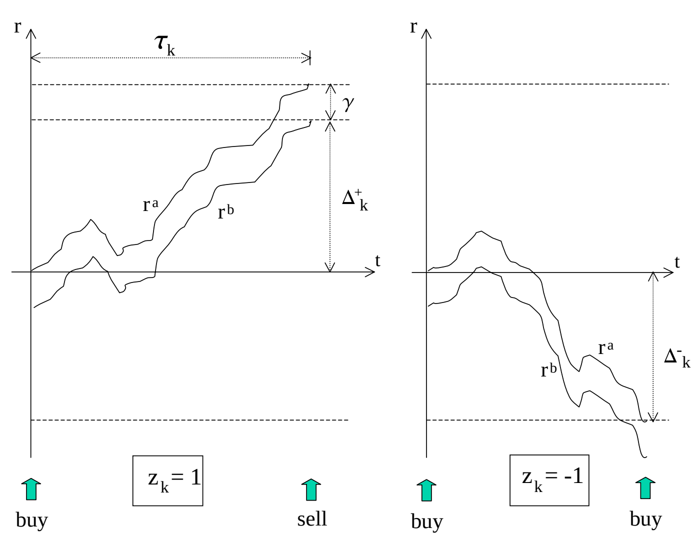

Fig. 1. We show (ask and bid) log-prices between the kth and the $(k+1)$th trade if the kth trade is an ask. We consider the case of a sell after a positive fluctuation ($z_k = 1$) of prices and a buy after a negative one ($z_k = -1$). As explained in the text this is the strategy followed by an investor in a repeated game.

ask price, but he has to accept the (lower) bid price. Instead no additional costs are involved if an ask follows an ask, since the investor is buying at a price equal to the value of the assets already in his portfolio. In that case the kth trade is a bid, one can reason in a similar way.

The decision to rebalance the portfolio can be due only to the size of the change in asset *value*. In fact the cost to modify is proportional to the *value*, so the best strategy is to rebalance only when the *value* changes enough to justify the costs incurred in the operation. An investor, who has no time constraints and who will repeat a bet again exactly in the same assets, does not decide to change his position on the basis of the time waited since the last trade (as in [12]). In fact between two portfolio rebalancing prices can remain unchanged or vary considerably: the costs paid are almost the same since they are proportional to asset value, even if in the first case the trade is useless. The lag length between two particular trades does not play any role in an optimal decision where no time constraints are externally imposed.

The investor can determine, on the basis of a rational criterion, the Δs at (or immediately after) the kth trade. Since no new information will arrive between the two changes (price process is stationary), there is no reason for him to change his mind afterwards or sell or buy before he has decided.[d] Furthermore log-price increments are identically independently distributed (hereafter i.i.d.) and hence every decision at the kth trade is completely independent from the past. It is made on the basis of his actual situation, i.e. on the kind of trade (buy or sell) he is performing.[e]

A trading strategy is just the repetition over the time of exactly the same game played at t_k. It is given by the two classes of parameters:

- $\{l\}$ which specifies the fraction of the capital invested on assets,
- $\{\Delta\}$ which is connected to the timing of the strategy,

where each class answers one of the two main questions of active portfolio managing we have dealt with in the introduction.

In this paper we select these parameters following the repeated game approach introduced by Kelly [13]. This approach has direct implications on the possible choices of the optimal parameters.

For example, a consequence of a repeated game in an i.i.d. financial market is that the fraction l and the Δs are always the same in the absence of costs. This

[d] We are considering a financial world where prices follow a given stochastic process not related to the behavior of a single investor. In this case it is equivalent to submit two limit orders immediately after the kth trade has taken place or to place a market order when the fluctuation of the given size is reached. In fact there is no market maker able to take advantage of the information in his order book and then there are no additional costs for the investor who shows his intentions just after the kth trade.

[e] To be more precise the trader's behavior depends also on his actual total wealth. We are assuming that he performs the same diversification between the two assets whether he invests \$10 or \$1000. One can always think that this is a reasonable approach, at least while his capital is within two boundary values.

implies that at the $(k+1)$th trade in order to obtain again the same fraction of capital invested in the risky asset, the investor has to buy shares when the price falls down and sell in the other case. If transaction costs are present both $\{l\}$ and $\{\Delta\}$ depend on the kind of trade (bid or ask) the investor performs at time t_k. In this case the trader still tries (as he would do in absence of costs) to buy at the "lower" price (i.e. after a negative return) and sell at the "higher" (i.e. after a positive return). This is a consequence of the fact that he repeats over and over again the same game and at each trade the only way he has to earn money is to sell at a price higher than the one he has bought. This condition on the trading rule is equivalent to requiring that before an ask (bid) is made the wealth owned in stocks should be smaller (larger) or equal than its value immediately after

$$\begin{array}{lcl} \exp(-\Delta_k^-) l_k W_k \leq l_{k+1} W_{k+1} & \multirow{2}{*}{if} & \text{he buys} \\ \exp(\Delta_k^+) l_k W_k \geq l_{k+1} W_{k+1} & & \text{he sells}\,. \end{array} \tag{1}$$

Following Kelly [13] we consider an investor who desires to maximize the growth rate of his capital W

$$\lambda \equiv \lim_{\Omega \to \infty} \frac{1}{\Omega} \ln \frac{W_{(\Omega)}}{W_{(0)}} \,. \tag{2}$$

It seems to be the most natural approach if the trader has an infinite horizon since capital will *almost surely* grow exponentially with a rate λ in the long run. If we limit our attention to the trading times we can rewrite the growth rate of the capital

$$\lambda = \frac{h}{\mathcal{T}} \,, \tag{3}$$

where we define

$$h \equiv \lim_{\mathcal{K} \to \infty} \frac{1}{\mathcal{K}} \sum_{k=0}^{\mathcal{K}} \ln \frac{W_{k+1}}{W_k} \,, \tag{4}$$

$\mathcal{K}$ is the number of trades up to the time Ω and

$$\mathcal{T} \equiv \lim_{\mathcal{K} \to \infty} \frac{1}{\mathcal{K}} \sum_{k=1}^{\mathcal{K}} \mathcal{T}_k \,, \tag{5}$$

is the mean exit time. In the case ergodic processes are considered, one can substitute the trading average in Eqs. (4) and (5) with the expectation value of the associate process.

We notice that this analysis, where prices are considered only when they hit the barriers, resembles Poincaré maps in dynamical systems. Going from continuous to discrete time setting, the quantities of dynamical interest are simply rescaled by the mean exit time $\mathcal{T}$ [14].

The discrete time framework is important in finance not only for a trader (who has to operate after finite lags). It is clearly relevant also at the level of the analysis of the financial market, because it allows us to show the connection of the capital growth rate, obtained via an optimal control policy, with quantities strictly related to Shannon entropy [15]. The main advantage is that this quantity can be measured on a historical dataset of financial assets or computed via elementary probability theory in simple cases. Furthermore this new framework will allow us to map the problem into one without costs.

3. Absence of Costs

Using the notation introduced above we summarize in this section the case of portfolio selection in the absence of transaction costs.

In this paper we model (ask and bid) log-prices with

$$dr_t = \mu dt + \sigma dw_t \,, \tag{6}$$

where μ is the drift and w_t a Wiener process with unitary variance.

We can safely assume a null risk-free interest rate R. The case $R > 0$ can always be recovered by simply considering the drift with respect to the interest rate (replacing μ with $\mu - R$) and adding R to the growth rate of the capital (2).

The capital at the $(k+1)$th trade is

$$W_{k+1} = [1 - l + l \exp(z_k \Delta^{z_k})] W_k \,. \tag{7}$$

As we have stressed in the previous section, the investor repeats exactly the same game over the time and then chooses the fractions l and the barriers $\{\Delta\}$ independently from the trade k considered. It is useful to observe that the capital (7) is a multiplicative process and then in the long run the growth rate is *almost surely*

$$\lambda(l; \Delta^+, \Delta^-) = \frac{p \ln[1 + l(\exp(-\Delta^-) - 1)] + (1-p) \ln[1 + l(\exp(\Delta^+) - 1)]}{\mathcal{T}} \,, \tag{8}$$

where $p = \pi(\Delta^+, \Delta^-)$ is the probability (A.1) to exit from the lower barrier and $\mathcal{T} = \tau(\Delta^+, \Delta^-)$ is the average exit time (A.2). Both quantities are derived in the appendix for the Wiener process (6).

We then divide the optimization problem in two steps, finding first the optimal fraction l at fixed barriers $\{\Delta\}$ and then searching the optimal Δs.

The optimal l satisfies

$$l(\Delta^+, \Delta^-) = \frac{1-p}{1 - \exp(-\Delta^-)} - \frac{p}{\exp(\Delta^+) - 1} \,. \tag{9}$$

Substituting the optimal value of l (9) into Eq. (8) we can write the numerator of the growth rate

$$h = p \ln \frac{p}{q} + (1-p) \ln \frac{1-p}{1-q} \tag{10}$$

as the Kullback entropy [16] between the probabilities p and q, where $q = q(\Delta^+, \Delta^-)$ is the martingale probability (A.3), which is the value of p for which no earning is possible. The relevant quantity is the information *available* to the investor: the

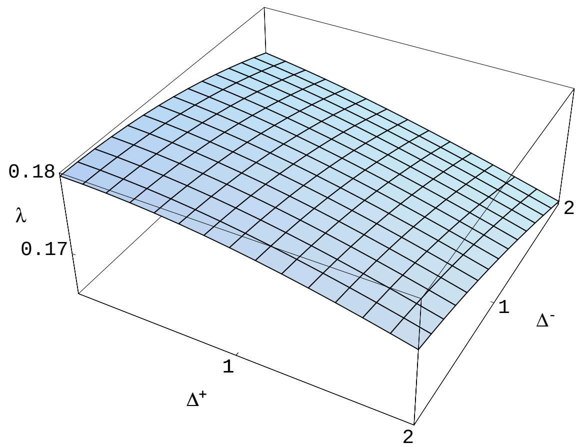

Fig. 2. Capital growth rate in the absence of costs for the optimal choice of l as a function of Δ^+ and Δ^-. The parameters chosen are $\mu = 0.1$ and $\sigma = 1$.

Kullback entropy per unit time w.r.t. the no-earning situation, i.e. roughly speaking the "distance" from the perfect situation of a fair game modeled by the probability (A.3).[f] In the absence and, as we shall show, in the presence of transaction costs, the *available* information has a clear financial interpretation for finite Δs. It is the maximum growth rate one can obtain having decided to trade after $\{\Delta\}$ fluctuations. The optimal trading rule does not depend any more on the control of the investor on his portfolio selling or buying stocks but it is completely related to the *available* information, a property of the financial asset considered as a stochastic signal. This quantity can be measured directly on the sequence of bid and ask prices of a financial asset.

Once we have selected the optimal l we can then choose the value of the barriers Δs that optimize the *available* information. In finance is not obvious that the maximum is reached for Δs arbitrarily small, because a financial asset is characterized by an "endogenous" finite scale due to the costs involved in the trading activity and to an "indeterminacy" of prices. These facts have been shown recently in the case of the Deutschemark/US Dollar exchange [18] and [19].

In Fig. 2 we plot the capital growth rate $\lambda(\Delta^+, \Delta^-)$ for the optimal choice of l. We notice that it is a non-increasing function of its arguments. The maximum is reached for $(\Delta^+, \Delta^- = 0)$, obtaining the well-known result that the optimal policy is continuous. It is an intuitive result due to the fact that a change in the portfolio does not cost anything and then the best solution is to use this *free* opportunity to continuously rebalance the investment.

[f] The *available* information is related to the Kolmogorov ϵ-entropy [17] (or with our notation the Δ-entropy) of the log-price process. In the particular case $q = 1/2$ considered by Kelly [13] h is (except for an additive constant) the entropy [15] of the dichotomic process $\{z\}$ and the entropy per unit time is the Kolmogorov Δ-entropy of the log-price.

We also observe that the maximum $(\Delta^+, \Delta^- = 0)$, is the only point where the gradient of the growth rate is zero. Therefore an investor, who changes his position at finite (but small) Δs, commits an error in the growth rate of the second order in Δ.

Performing the limit $\Delta^{+,-} \to 0$ in the equations (9) and (10) one obtains the optimal capital fraction

$$l^* = \frac{\mu}{\sigma^2} + \frac{1}{2} \tag{11}$$

and the optimal growth rate

$$\lambda^* = \frac{\sigma^2}{2}\, l^{*^2}\,. \tag{12}$$

We do not allow the trader to borrow money from a bank or short selling of stock. An optimal portfolio suggests keeping the same fraction forever (l^* does not depend on k) and, because we are dealing with a portfolio including stocks as risky assets, a never-ending borrowing or short selling position does not appear realistic. The considered cases correspond to

$$-1 \leq \frac{2\mu}{\sigma^2} \leq 1\,. \tag{13}$$

If this ratio is 1, the best solution is to transfer all the money to the stocks; instead the opposite limit corresponds to having no money in the risky asset. In the following we consider only transaction costs in the case with drift μ and variance σ which satisfy condition (13).

4. Solution in the Presence of Costs

In the portfolio selection in absence of costs there is a unique asset price. It is natural to associate the asset value with this price. If transaction costs are present the asset value is a convention instead. One obtains the bid price selling the asset and buys it at the ask price. The value can be any price between these two. As we have discussed in Sec. 2 a point of view which seems natural is to consider a value *relative* to the investor's last trade: it is the bid price when he sells and the ask price when he buys.

Let us focus our attention again on the period immediately after the kth trade. We are considering a simple case of a no-memory process. The only piece of information the investor has to remember of his past is his previous trade, i.e. if he has bought or sold assets.

The initial (final) *state* at the k $(k+1)$ trade is defined as

$$\xi, \eta = \begin{cases} a\ (ask) & \text{when the trader} \quad \text{buys} \\ b\ \ (bid) & \phantom{\text{when the trader}} \quad \text{sells}\,. \end{cases}$$

The capital after the $(k+1)$th trade is

$$W_{k+1} = [1 + l_\xi(\exp(z_k \Delta_\xi^{z_k}) - 1)]W_k \equiv \omega_{\xi\eta} W_k \,, \tag{14}$$

where z_k has been defined in Sec. 2 as the sign of the return between k and $k+1$ and

$$\eta = a \quad \text{when} \quad z_k = -1$$

$$\eta = b \quad \text{when} \quad z_k = +1 \,,$$

i.e. in a repeated game the investor buys after a negative fluctuation and sells after a positive.

The problem is then completely defined by a Markovian transition matrix between the initial and final *state*

$$V = \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix} , \tag{15}$$

where

$$\begin{cases} \alpha = 1 - \pi(\Delta_a^+ + \gamma, \Delta_a^-) \\ \beta = \pi(\Delta_b^+, \Delta_b^- + \gamma) \end{cases}$$

and $\pi(\Delta^+, \Delta^-)$ is the probability (A.1).

The probabilities of the *states* a and b satisfy the following relation

$$\frac{p_a}{p_b} = \frac{V_{ba}}{V_{ab}} = \frac{\beta}{\alpha} \,.$$

We have mapped the original portfolio selection into one where no costs are present but a Markovian memory must be considered. The capital growth rate for an investor who follows this strategy is:

$$\lambda = \frac{\sum_{\xi,\eta} p_\xi V_{\xi\eta} \ln \omega_{\xi\eta}}{\mathcal{T}} \,, \tag{16}$$

where the average exit time $\mathcal{T}$ is

$$\mathcal{T} = \sum_\xi p_\xi \mathcal{T}_\xi$$

with

$$\mathcal{T}_a = \tau(\Delta_a^+ + \gamma, \Delta_a^-)$$

$$\mathcal{T}_b = \tau(\Delta_b^+, \Delta_b^- + \gamma) \,.$$

Equation (16) is the core of the paper: the selection of the optimal trading rule from the prospective of the *relative* value leads us to reduce our task to an optimization problem with a finite-state Markovian chain. The problem can then be faced with elementary probability theory in discrete spaces.

In the following we shall find the $\{l\}$ and $\{\Delta\}$ which optimize the growth rate (16). We consider then an ansatz, with l and $\{\Delta^+, \Delta^-\}$ independent from the initial *state*, as in the case of an absence of costs. We discuss in detail this approximation underlining why it can be relevant for practical purposes.

4.1. *Exact solution*

We have shown how the portfolio selection in presence of transaction costs has been mapped into a Markovian problem in the absence of costs. We find here the optimal solution in the same way as in the previous section: maximizing $\{l\}$ first and then finding the optimal barriers.

The two optimal values of l are:

$$\begin{aligned} l_a(\Delta_a^+, \Delta_a^-) &= \frac{\alpha}{1-\exp(-\Delta_a^-)} - \frac{1-\alpha}{\exp(\Delta_a^+)-1} \\ l_b(\Delta_b^+, \Delta_b^-) &= \frac{1-\beta}{1-\exp(-\Delta_b^-)} - \frac{\beta}{\exp(\Delta_b^+)-1}\,. \end{aligned} \tag{17}$$

Substituting l_a and l_b in (16) one obtains

$$\lambda = \frac{\sum_{\xi,\eta} p_\xi V_{\xi\eta} \ln \frac{V_{\xi\eta}}{Q_{\xi\eta}}}{\mathcal{T}}\,, \tag{18}$$

where

$$Q = \begin{pmatrix} 1-q_a & q_a \\ q_b & 1-q_b \end{pmatrix},$$

with

$$\begin{cases} q_a = 1 - q(\Delta_a^+, \Delta_a^-) \\ q_b = q(\Delta_b^+, \Delta_b^-)\,. \end{cases}$$

We notice that the numerator of Eq. (18) is, for a Markov chain, the quantity equivalent to the Kullback entropy (10) obtained in the no memory case. The right side of Eq. (18) is the *available* information in the Markovian chain case, a measurable quantity of the underlying process even if transaction costs are present.

The investor tries then to select the Δs that maximize the *available* information (18). We have observed in Fig. 1 that the transaction costs play a role only when the trader changes his *state*: every modification of his portfolio still remaining in the same *state* costs nothing. Of course he uses this *free* opportunity to rebalance his portfolio as often as he can, i.e.

$$\Delta_a^-, \Delta_b^+ \to 0\,.$$

One can show that capital growth rate $\lambda(\Delta_a^+, \Delta_b^-)$ is a non-increasing function of both Δ_a^+ and Δ_b^-. Because of this monotonicity of the *available* information the maximum values are on the boundaries of the allowed zone for Δs.

We remind ourselves that we are looking for a solution where the trader sells at a higher price and buys at a lower, and then we have to check that the Δs satisfy condition in (1), which are equivalent to:

$$\begin{cases} \nu_a(\Delta_a^+) \equiv \dfrac{e^{\Delta_a^+}\; l_a(\Delta_a^+)}{1+(e^{\Delta_a^+}-1)\; l_a(\Delta_a^+)} \geq l_b(\Delta_b^-) \\[2ex] \nu_b(\Delta_b^-) \equiv \dfrac{e^{-\Delta_b^-}\; l_b(\Delta_b^-)}{1+(e^{-\Delta_b^-}-1)\; l_b(\Delta_b^-)} \leq l_a(\Delta_a^+)\,. \end{cases} \tag{19}$$

It is easy to verify that ν_ξ values are monotonous and in particular ν_a is a strictly increasing function of its argument and ν_b strictly decreasing. Inverting (19) one obtains:

$$\begin{cases} \Delta_a^+ \geq \nu_a^{-1}[l_b(\Delta_b^-)] \\ \Delta_b^- \geq \nu_b^{-1}[l_a(\Delta_a^+)]\,. \end{cases} \tag{20}$$

In Fig. 3 we show the region of Δs allowed by condition (20). We observe in particular that $\nu_a^{-1}[l_b(\Delta_b^-)]$ and $\nu_b^{-1}[l_a(\Delta_a^+)]$ cross at the point $\Delta_a^+ = \Delta_b^- = \Delta$ where Δ is given by equation

$$\frac{2\mu}{\sigma^2}\sinh\left(\frac{\Delta}{2}\right) = \sinh\frac{\mu}{\sigma^2}(\Delta+\gamma)\,. \tag{21}$$

We notice that it is possible to show after some algebra that the tangents of the two curves are parallel to the axes in the crossing point $(\Delta_a^+, \Delta_b^- = \Delta)$ and then Δ is the lowest allowed value for both Δ_a^+ and Δ_b^-. Since the capital growth rate is a non-increasing function of the two barriers, Eq. (21) is the optimal choice for the barriers of the portfolio selection problem in presence of transaction costs.

In Fig. 4 we plot the optimal values of Δ as a function of $2\mu/\sigma^2$ obtained from Eq. (21) for a particular choice of γ observing that Δ goes to infinity when $2\mu/\sigma^2$ approaches the limit values of -1 or 1.

One can also show that $(\Delta_a^-, \Delta_b^+ = 0; \Delta_a^+, \Delta_b^- = \Delta)$ is the only point where the gradient of the growth rate (18) is the null vector. This fact is particularly relevant

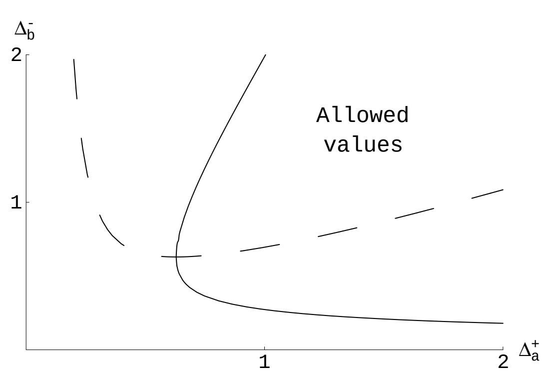

Fig. 3. The values of Δs allowed by conditions (20) for $\mu/\sigma^2 = 0.1$ and $\gamma = 0.01$. The full line represents $\nu_a^{-1}[l_b(\Delta_b^-)]$ and the dashed $\nu_b^{-1}[l_a(\Delta_a^+)]$.

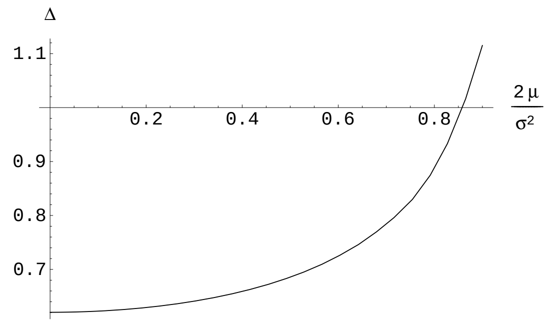

Fig. 4. Optimal values of Δ *vs* $2\mu/\sigma^2$ for $\gamma = 0.01$. We plot only the positive values because Δ is an even function of μ.

for a trader, because the optimal solution requires on average to rebalance the portfolio after a time equal to zero (a consequence of $\Delta_a^-, \Delta_b^+ \to 0$). A null gradient of the capital growth rate on the optimal solution implies, as in the absence of costs case, that a suboptimal choice of the barriers causes a small error in the capital growth rate.

The optimal fractions (17) become

$$\begin{aligned} l_a(\Delta) &= \frac{2\mu/\sigma^2}{1-\exp(-\frac{2\mu}{\sigma^2}(\Delta+\gamma))} - \frac{1}{\exp(\Delta)-1} \\ l_b(\Delta) &= \frac{1}{1-\exp(-\Delta)} - \frac{2\mu/\sigma^2}{\exp(\frac{2\mu}{\sigma^2}(\Delta+\gamma))-1}, \end{aligned} \tag{22}$$

where Δ is given by Eq. (21). We observe that

$$l_a(\Delta) = 2l^* - l_b(\Delta),$$

i.e. l^* is the mean value of $l_a(\Delta)$ and $l_b(\Delta)$.

In Fig. 5 we show the values of l_ξ (22) for two different values of γ and the fraction l^* of the capital in the no-cost case (9).

The optimal growth rate of the portfolio in presence of transaction costs is

$$\lambda_O = \frac{\sigma^2}{2} l_a(\Delta) l_b(\Delta) \leq \lambda^*, \tag{23}$$

where λ^* is the one obtained in the no-costs case (12). We notice that the optimal growth rate of the capital is proportional to the square of the geometric average of l_a and l_b and λ^* to the square of their arithmetic average. The geometric average is always lower or equal to the arithmetic average l^* and equal only when $l_a = l_b$, i.e. for $2\mu/\sigma^2 = -1, 1$. This fact has a simple interpretation: these are the cases where the trader maintains his position forever; in these situations no costs are paid, except at least once when the investment begins.

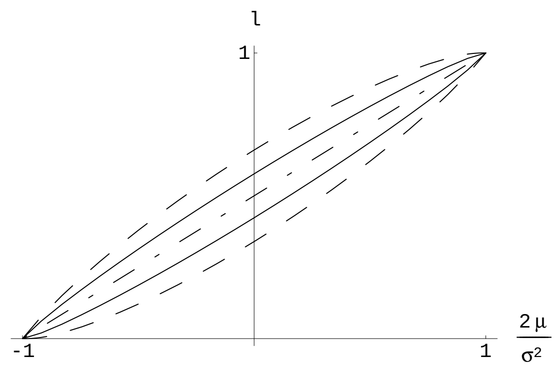

Fig. 5. Optimal values of l_a and l_b *vs* $2\mu/\sigma^2$. The fraction of the capital l_a is greater than the value l^* obtained in the absence of costs (dot dashed line) while l_b is lower. The dashed line corresponds to the value of $\gamma = 0.1$ and the full line to $\gamma = 0.01$.

4.2. *An approximate solution*

The optimal solution implies that the average time after which the trader should modify his position is zero. In this subsection we suggest a feasible approximation of the optimal trading rule, considering the simplest ansatz for the investor's strategy. We choose both the barriers and the fraction independent on the initial *state* ξ. This should be a reasonable approximation, as explained in the no-costs case, because we are dealing with a no memory process and a growth rate maximizer. We notice that conditions in (1) are automatically satisfied once we choose a time-independent value of l.

The capital growth rate (16) becomes

$$\lambda = \frac{p\ln[1 + l(\exp(-\Delta^-) - 1)] + (1-p)\ln[1 + l(\exp(\Delta^+) - 1)]}{p\mathcal{T}_a + (1-p)\mathcal{T}_b}\,, \tag{24}$$

where $p = \beta/(\alpha + \beta)$.

We have mapped the problem into one in the absence of costs, as the one considered in Sec. 3 with probability p to exit from the lower barrier and average exit time $\mathcal{T} = p\mathcal{T}_a + (1-p)\mathcal{T}_b$. The optimal value of l is given by Eq. (9). Substituting this value into Eq. (24), one obtains again that the growth rate can be written as the ratio between the Kullback entropy (10) and the average exit time $\mathcal{T}$.

In Fig. 6 we plot the capital growth rate (3) as a function of the barriers Δ^+ and Δ^- calculated for the optimal l (9). We observe that the maximum is reached for finite values of the barriers. Thus the investor (on average) changes his portfolio after a finite time, i.e. he follows a discrete time trading rule.

In Fig. 7 we plot the values of the barriers for which the optimal growth rate is reached. The fraction l of the capital computed on the optimal barriers is plotted

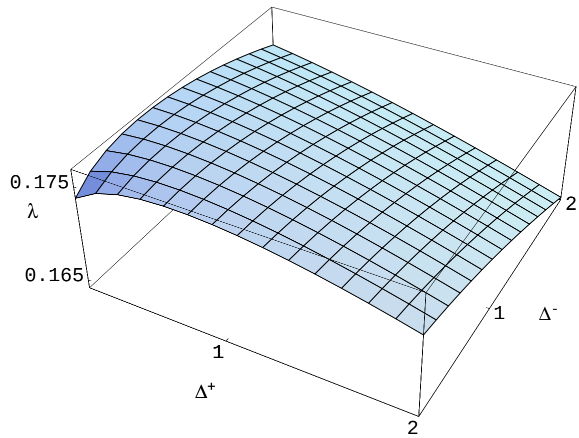

Fig. 6. Capital growth rate for the optimal choice of l as a function of Δ^+ and Δ^-. The parameters are $\mu = 0.1$, $\sigma = 1$, $\gamma = 0.01$.

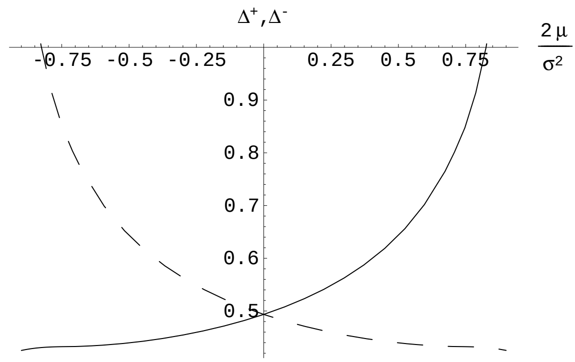

Fig. 7. Optimal values of Δ^- (full line) and Δ^+ (dashed line) *vs* $2\mu/\sigma^2$ for $\gamma = 0.01$.

in Fig. 8. We notice that

$$\begin{aligned} \Delta^-(-\mu) &= \Delta^+(\mu) \\ \Delta^+(-\mu) &= \Delta^-(\mu) \\ l(-\mu) &= 1 - l(\mu) \end{aligned}$$

as a consequence of the symmetry

$$\lambda_{-\mu}(l; \Delta^+, \Delta^-) = \lambda_\mu(1 - l; \Delta^-, \Delta^+) - \mu \tag{25}$$

of the capital growth rate (24).

In Fig. 9 we compare the optimal capital growth rate with the approximate one. We notice that even for so large a (and unrealistic) transaction cost the differences between the two are negligible for $\mu \geq 0$, which is the typical value for stocks,

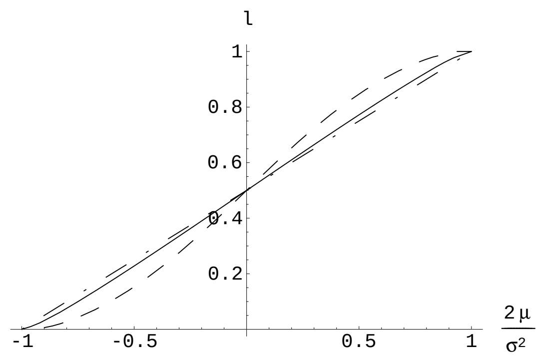

Fig. 8. Optimal values of l *vs* $2\mu/\sigma^2$. The dashed line corresponds to the value of $\gamma = 0.1$ and the full line to $\gamma = 0.01$. We plot also the value of l in the absence of costs (dot dashed line).

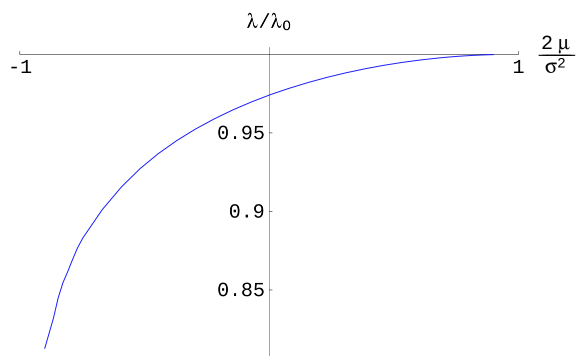

Fig. 9. Capital growth rate of the approximate solution λ rescaled with the exact one λ_O as a function of $2\mu/\sigma^2$ for $\gamma = 0.1$.

since log-price drift is usually greater then the risk-free interest rate. As we would expect, it is reasonable to limit oneself to a time-independent trading rule if returns are i.i.d.

5. Conclusions

We consider in this paper a new way of treating the transaction costs problem: the investor modifies his position only when the asset *value* changes of a "relevant" amount and this *value* depends on the last trade (bid or ask).

This point of view presents several advantages compared with the "traditional" one. We split the optimal trading rule in two main questions for an investor and we show that it allows us to map the portfolio selection problem with transaction costs into a Markovian one without costs. We connect the quantities of interest (growth rate and optimal portfolio strategy) to the *available* information which depends on average exit times and probabilities to reach a barrier. If log-prices are modeled by a Wiener process, as assumed in this paper, both quantities can be computed using elementary probability theory, allowing analytical formulas for both the strategy (the fraction l of wealth invested in the risky asset and the barriers Δs) and the *almost sure* capital growth rate of the optimal solution.[g]

We show that the exact solution of the problem in the presence of transaction costs breaks the time invariance of the investment. However such a strategy is not feasible in practice because the trader should modify his portfolio after a time which is on average zero. In Subsec. 4.2 we considered a suboptimal strategy feasible for a trader who wants to use it. This strategy, where the broken symmetry is restored, shows small differences in the capital growth rate with respect to the optimal one and involves only a finite number of transactions in finite time.

We have assumed in this paper that log-price is a Wiener process (i.e. the process is continuous and increments are i.i.d.) to understand and illustrate the approach in the simplest situation. In this case at each trade the strategy depends on a dichotomic variable (related to the two possibilities of the trader: buy or sell) that we have called *state*. The main advantage of the proposed approach is that we focus only on the times the trader modifies his position, avoiding a continuous following of price changes, allowing us to have a general methodology[h] connected to the way a financial decision is taken in practice.

In the general case, price process presents jumps and then the log-price does not arrive exactly on the barrier but generally overcomes it: in this case the *state*, instead of assuming only two values, has many values depending on the distance of the log-price from the barrier.

Returns also have a Markovian memory. It is intuitive that events far in the past are irrelevant in a portfolio selection. Portfolio strategy thus depends on the last m trades (instead of only the last one as in the Wiener case), where the Markovian order m of the process can be determined with standard techniques of the information theory [20].

Even the case of non-constant transaction costs is not a problem: we have to consider the ask price when investor buys and the bid price when it sells.

This approach then allows a straightforward generalization for the case with memory and jumps; it is then particularly relevant in the case of a real market, where exit times and probabilities to reach (and overcome) a barrier can be measured directly on a historical dataset [21].

In the general case of a price driven by a stationary stochastic process, the selection of an optimal portfolio allocation in the presence of transaction costs depends on the *available* information of the asset price: a quantity that can be measured on real data or computed analytically if an elementary price model is assumed. The Wiener process can be considered a toy model for log-prices. Nevertheless for

[g]In the case the trader is not only interested in the behavior of the capital growth rate in the long run but also in controlling the short term fluctuation around the limit value, the same technique can be easily extended to all the utilities in the HARA class.

[h]Under the hypothesis of a stationary price process.

a trader who waits until relevant price changes appear, this model captures the essential features of the portfolio selection, leading also to a deeper understanding of the role of approximate but feasible strategies.

Appendix A

In this appendix we compute the probabilities and the average exit times of a Wiener process as limit of a random walk on a one-dimensional lattice; in this case both quantities can be obtained with the elementary probability theory (see for example Chap. 14 of [11]).

The walker goes after a time step ϵ to the right of a lattice step $\sigma\sqrt{\epsilon}$ with probability

$$p_\epsilon = \frac{1}{2}\left(1 + \frac{\mu}{\sigma}\sqrt{\epsilon}\right)$$

and to the left with probability $1 - p_\epsilon$, where μ and σ are the parameters of the Wiener process (6).

Starting from zero the probability to reach a barrier situated at $-\Delta^-$ before hitting a barrier at Δ^+ is

$$\pi_\epsilon(\Delta^+, \Delta^-) = \begin{cases} \dfrac{\rho_\epsilon^{\Delta^+} - 1}{\rho_\epsilon^{\Delta^+ + \Delta^-} - 1} & \text{if } \mu \neq 0 \\ \dfrac{\Delta^+}{\Delta^+ + \Delta^-} & \text{if } \mu = 0 \end{cases}$$

where

$$\rho_\epsilon = \left(\frac{p_\epsilon}{1 - p_\epsilon}\right)^{\frac{1}{\sigma\sqrt{\epsilon}}} .$$

The average time to exit from one of the two barriers is

$$\tau_\epsilon(\Delta^+, \Delta^-) = \begin{cases} \dfrac{1}{\mu}[\Delta^+ - (\Delta^+ + \Delta^-)\pi_\epsilon] & \text{if } \mu \neq 0 \\ \dfrac{1}{\sigma^2}\Delta^+\Delta^- & \text{if } \mu = 0 \, . \end{cases}$$

Performing the limit $\epsilon \to 0$, one recovers the Wiener process defined in (6) and defining

$$\rho \equiv \lim_{\epsilon \to 0} \rho_\epsilon = \exp\left[\frac{2\mu}{\sigma^2}\right] ,$$

we obtain the quantities of interest. The probability to hit the lower barrier is

$$\pi(\Delta^+, \Delta^-) = \begin{cases} \dfrac{\rho^{\Delta^+} - 1}{\rho^{\Delta^+ + \Delta^-} - 1} & \text{if } \mu \neq 0 \\ \dfrac{\Delta^+}{\Delta^+ + \Delta^-} & \text{if } \mu = 0 \, , \end{cases} \tag{A.1}$$

and the average exit time is

$$\tau(\Delta^+, \Delta^-) = \begin{cases} \dfrac{1}{\mu}\left[\Delta^+ - (\Delta^+ + \Delta^-)\pi\right] & \text{if } \mu \neq 0 \\ \dfrac{1}{\sigma^2}\Delta^+\Delta^- & \text{if } \mu = 0\,. \end{cases} \tag{A.2}$$

We observe that both the probability (A.1) and the exit time (A.2) are continuous in μ.

We define martingale probability as

$$q(\Delta^+, \Delta^-) = \frac{1 - e^{-\Delta^+}}{1 - e^{-\Delta^+ - \Delta^-}}\,, \tag{A.3}$$

the one with respect to which asset *value* (on the barriers) is a martingale process, i.e.

$$qe^{-\Delta^-} + (1-q)e^{\Delta^+} = 1\,.$$

Acknowledgments

We thank the European Community TMR Grant "Financial Market Efficiency and Economic Efficiency". We thank also Markus Abel for a careful reading of the manuscript, Massimo Cencini for an enlightening conversation about Markovian processes, Filippo Cesi for a nice discussion on exit times in Wiener processes, Michele Pasquini, Maurizio Serva and Davide Vergni for several comments and suggestions on the subject, Angelo Vulpiani for a huge number of discussions about Kelly's strategies and my father for never ending encouragement all throughout the period I was putting together this work.

References

[1] J. Ingersoll, *Theory of Financial Decision Making*, Rowman & Littlefield (1987).
[2] D. Duffie, *Dynamical Asset Pricing Theory*, Princeton University Press (1992).
[3] R. Merton, *Continuous-Time Finance*, Blackwell (1990).
[4] M. Taksar, M. J. Klass and D. Assaf, *A diffusion model for optimal portfolio selection in presence of brokerage fees*, Math. of Oper. Research **13** (1988) 277–294.
[5] G. M. Constantinides, *Multiperiod consumption and investment behavior with convex transactions costs*, Management Science **25** (1991) 1127–1137.
[6] B. Dumas and E. Luciano, *An exact solution to a dynamic portfolio choice problem under transaction costs*, J. Finance **46** (1991) 577–595.
[7] M. Akian, J. L. Menaldi and A. Sulem, *On an investment-consumption model with transaction costs*, SIAM J. Control and Optim. **34** (1996) 329–364.
[8] R. C. Merton, *Optimum consumption and portfolio rules in a continuous-time model*, J. Econ. Theory **3** (1971) 373–413.
[9] J. M. Harrison, *Brownian Motion and Stochastic Flow System*, John Wiley (1991).
[10] I. Karatzas and S. E. Shreve, *Brownian Motion and Stochastic Calculus*, Springer-Verlag (1985).

[11] W. Feller, *An Introduction to Probability Theory and its Applications I*, John Wiley, New York (1971).
[12] M. Serva, *Optimal lag in dynamical investment*, Int. J. Theoretical and Applied Finance, to appear.
[13] Jr. J. L. Kelly, *A new interpretation of the information rate*, Bell Syst. Tech. J. **35** (1956) 917–926.
[14] R. Badii and A. Politi, *Complexity: Hierarchical Structures and Scaling in Physics*, Cambridge University Press, Cambridge (1997).
[15] C. E. Shannon, *The mathematical theory of communication*, Bell Syst. Tech. J. **27** (1948) 379–423, 623–656.
[16] S. Kullback and R. A. Leibler. *On information and sufficiency*, Ann. Math. Statist. **27** (1956) 79–86.
[17] A. N. Kolmogorov, *On the Shannon theory of information transmission in the case of continuous signals*, IRE Trans. on Inform. Theory **IT-2** (1956) 102–108.
[18] R. Baviera, M. Parquini, M. Serva, D. Vergni and A. Vulpioni, *Efficiency in foreign exchange markets*, available in http://papers.ssrn.com (1999).
[19] M. Pasquini and M. Serva, *Indeterminacy in foreign exchange markets*, Physica **A277** (2000) 228–238.
[20] R. Baviera, D. Vergni and A. Vulpiani, *Markovian approximation in foreign exchange markets*, Physica **A280** (2000) 566–581.
[21] R. Baviera, *Weak efficiency in absence and in presence of transaction costs*, in preparation.