# SELF-FINANCING MARKETS AND EVENTUAL ARBITRAGE

J. F. CARRIÈRE
Mathematical Sciences
University of Alberta
Edmonton, Canada
j.carriere@ualberta.ca

*Abstract*: In this article, we argue that the self-financing axiom with mild assumptions on the conditional expected returns yields a market with an eventual arbitrage. This is accomplished by minimizing the conditional variance of a trade when the conditional expectation is a fixed constant. Examples with common processes shows that most models yield an eventual arbitrage. As a final application, a cost model is applied to prices of stripped coupon and principal payments on U.S. government bonds, where rates of return are estimated and lower bounds on market costs given.

*Key Words*: Multivariate stochastic processes in discrete time, trading strategies, STRIPS.

# 1 Introduction

In this article, we discuss the notion of an eventual arbitrage and give examples of how they can be constructed. Our main thesis is that an observable opportunity to exercise an eventual arbitrage is actually non-exploitable because of costs. Thus, estimates of expected gains can be interpreted as a lower bound on market costs.

In the next section, we introduce the notion of a self-financing maket where costs are allowed. Next, we show that variance minimizing strategies lead to an eventual arbitrage. Using price data on STRIPS, we calibrate the model and construct a good predictive strategy. Using our day trader's cost model, we get an estimate on market costs.

# 2 Self-Financing Markets

Consider a self-financing market with a finite collection of $p$ assets where the market prices at time $t$ are denoted by the column vector $\boldsymbol{P}_t = [P_{t1}, \dots, P_{tp}]'$. Assume that $\boldsymbol{P}_t$ is a stochastic process adapted to a history ($\sigma$-field) denoted as $\mathcal{H}_t$. For our purposes, let $\mathcal{P}$ denote the physical probability function induced by the price process $\boldsymbol{P}_t$ with the collection of all measurable events being $\mathcal{H}_\infty$. Also let E denote the associated expectation operator. Next, let $\boldsymbol{\phi} = [\boldsymbol{\phi}_0, \boldsymbol{\phi}_1, \dots]'$ denote a trading strategy where the trading position at time $t = 0, 1, \dots$ is $\boldsymbol{\phi}_t$. We assume that $\boldsymbol{\phi}_t = [\phi_{t1}, \dots, \phi_{tp}]'$ is a $p \times 1$ random vector that is measurable with respect to $\mathcal{H}_t$. Let $V_t(\boldsymbol{\phi})$ denote the value of a portfolio constructed of these $p$ assets according to a trading stategy $\boldsymbol{\phi}$. This is equal to

$$V_t(\boldsymbol{\phi}) = \boldsymbol{\phi}_t \cdot \boldsymbol{P}_t. \tag{2.1}$$

Next, let $D_t \geq 0$ denote an infusion of cash into the fund at time $t$ that is called the dividend and let $C_t \geq 0$ denote an outlay of cash from the fund at time $t$ that is called the cost.

**Definition 1.** We will say that the portfolio is *self-financing with cash-flow* whenever the value $V_t(\boldsymbol{\phi})$

has the representation

$$\boldsymbol{\phi}_t \cdot \boldsymbol{P}_t = \boldsymbol{\phi}_{t-1} \cdot \boldsymbol{P}_t + D_t - C_t. \tag{2.2}$$

If $D_t = 0$ and $C_t = 0$, then we will simply say that the portfolio is *self-financing*. In this special case $(\boldsymbol{\phi}_t - \boldsymbol{\phi}_{t-1}) \cdot \boldsymbol{P}_t = 0$. For more information about self-financing portfolios, consult Musiela and Rutkowski (1998).

Next, denote the change in prices by $\boldsymbol{\Delta}_t = \boldsymbol{P}_{t+1} - \boldsymbol{P}_t = [\Delta_{t1}, \ldots, \Delta_{tp}]'$. We find that

$$V_{t+1} - V_t = \boldsymbol{\phi}_t \cdot \boldsymbol{\Delta}_t + D_t - C_t. \tag{2.3}$$

Moreover, we find that

$$V_t(\boldsymbol{\phi}) = V_0(\boldsymbol{\phi}) + \sum_{k=0}^{t-1} [\boldsymbol{\phi}_k \cdot \boldsymbol{\Delta}_k + D_k - C_k]. \tag{2.4}$$

Suppose that $D_t = 0$ for all $t$ and that the cost function is equal to

$$C_t = |\boldsymbol{\phi}_t| \cdot \boldsymbol{c_t}, \tag{2.5}$$

where $\boldsymbol{c_t} = [c_{t1}, \ldots, c_{tp}]'$ is measurable with respect to $\mathcal{H}_t$ and where $c_{ti} > 0$ for $i \in \{1, 2, \ldots, p\}$, are the costs of holding one security. We call this the *day trader's* cost model. We now introduce the notion of *non-exploitable* markets.

**Definition 2.** Suppose that we have a self-financing portfolio with $D_t = 0$ and $C_t > 0$ for all $t$. We will say that the market at time $t$ is **non-exploitable** by using a trading strategy $\boldsymbol{\phi}_t$ whenever $\mathrm{E}[\boldsymbol{\phi}_t \cdot \boldsymbol{\Delta}_t - C_t | \mathcal{H}_t] \leq 0$.

As a special case, non-exploitability implies that using a day trader's cost model yields

$$\boldsymbol{\phi}_t \cdot \boldsymbol{\mu}_t \leq |\boldsymbol{\phi}_t| \cdot \boldsymbol{c_t}, \tag{2.6}$$

where the conditional expected change in prices is defined by

$$\boldsymbol{\mu}_t = \mathrm{E}[\boldsymbol{\Delta}_t | \mathcal{H}_t]. \tag{2.7}$$

## 3 Self-Financing Strategies That Minimize Variance

In this section, we apply a Markowitz model to portfolio optimization and find that this is sufficient for an eventual arbitrage. For more information about the Markowitz model, consult Panjer, Boyle, et al (1998). For the ensuing discussion, we assume that $D_t = C_t = 0$ and we define

$$\boldsymbol{\Sigma}_t = \mathrm{Var}[\boldsymbol{\Delta}_t | \mathcal{H}_t]. \tag{3.1}$$

Now, suppose that $\boldsymbol{\mu}_t = \mathbf{0}$ for all $t$ then $\boldsymbol{P}(t)$ is a martingale and $\mathrm{E}[V_t(\boldsymbol{\phi})] = V_0(\boldsymbol{\phi})$ for all $\boldsymbol{\phi}$. In this case, prediction is not possible. Next, let $\mu_{tl}$ denote the $l$-th coordinate of $\boldsymbol{\mu}_t$. Note that if $\mu_{tl} \neq 0$ for all $l$ and $t$ then we can set $\phi_{tl} = \frac{r_t/p}{\mu_{tl}}$ with $r_t \in \mathbb{R}$. Thus $\mathrm{E}[V_t(\boldsymbol{\phi})] = V_0(\boldsymbol{\phi}) + \sum_{k=0}^{t-1} r_k$ and so the expected value of a self-financing portfolio can be set at any level. The same cannot be said of the variance. In this paper, we will focus on strategies that minimize a variance for a fixed expectation. That is, we want to choose $\boldsymbol{\phi}_t$ so that $\mathrm{Var}[\boldsymbol{\phi}_t \cdot \boldsymbol{\Delta}_t | \mathcal{H}_t]$ is minimized subject to the constraint that $\mathrm{E}[\boldsymbol{\phi}_t \cdot \boldsymbol{\Delta}_t | \mathcal{H}_t] = r_t$, where $r_t > 0$. That is, we need to find $\boldsymbol{\phi}_t$ so that $\boldsymbol{\phi}_t{}' \boldsymbol{\Sigma}_t \boldsymbol{\phi}_t$ is minimized subject to the constraint that $\boldsymbol{\phi}_t{}' \boldsymbol{\mu}_t = r_t$. The well-known solution is

$$\boldsymbol{\phi}_t = \frac{r_t \boldsymbol{\Sigma}_t^{-1} \boldsymbol{\mu}_t}{\boldsymbol{\mu}_t{}' \boldsymbol{\Sigma}_t^{-1} \boldsymbol{\mu}_t}, \tag{3.2}$$

whenever, the inverse $\boldsymbol{\Sigma}_t^{-1}$ exists and $\boldsymbol{\mu}_t{}'\boldsymbol{\Sigma}_t^{-1}\boldsymbol{\mu}_t > 0$. As a special case, if $p = 1$ then $\boldsymbol{\phi}_t = \frac{r_t}{\boldsymbol{\mu}_t}$. In general, more constraints can be put on the strategy $\boldsymbol{\phi}_t$ but we will not be investigating the general problem. Now, consider the variance

$$\mathrm{Var}[\boldsymbol{\phi}_t \cdot \boldsymbol{\Delta}_t | \mathcal{H}_t] = \frac{r_t^2}{\boldsymbol{\mu}_t{}'\boldsymbol{\Sigma}_t^{-1}\boldsymbol{\mu}_t}. \tag{3.3}$$

Let us calculate the variance of $V_t$. This is equal to

$$\mathrm{Var}[V_t] = \sum_{k=0}^{t-1}\sum_{l=0}^{t-1} \mathrm{Cov}[\boldsymbol{\phi}_k \cdot \boldsymbol{\Delta}_k, \boldsymbol{\phi}_l \cdot \boldsymbol{\Delta}_l]. \tag{3.4}$$

If $k \neq l$ then $\mathrm{Cov}[\boldsymbol{\phi}_k \cdot \boldsymbol{\Delta}_k, \boldsymbol{\phi}_l \cdot \boldsymbol{\Delta}_l] = 0$ because $\boldsymbol{\phi}_k \cdot \boldsymbol{\mu}_k = r_k \in \mathbb{R}$ is non-stochastic, by construction. Note that no independent-increment assumption is made. In the case that $k = l$ we get $\mathrm{Cov}[\boldsymbol{\phi}_k \cdot \boldsymbol{\Delta}_k, \boldsymbol{\phi}_k \cdot \boldsymbol{\Delta}_k] = \mathrm{E}[\mathrm{Var}[\boldsymbol{\phi}_k \cdot \boldsymbol{\Delta}_k | \mathcal{H}_k]]$. Thus we get

$$\mathrm{Var}[V_t] = \sum_{k=0}^{t-1} r_k^2 \mathrm{E}\left[ (\boldsymbol{\mu}_k{}'\boldsymbol{\Sigma}_k^{-1}\boldsymbol{\mu}_k)^{-1} \right]. \tag{3.5}$$

In the constant variance model, we have

$$\gamma_k^2 \equiv \mathrm{E}\left[(\boldsymbol{\mu}_k{}'\boldsymbol{\Sigma}_k^{-1}\boldsymbol{\mu}_k)^{-1}\right] = \gamma^2 > 0, \tag{3.6}$$

for all $k$ and so $\mathrm{Var}[V_t] = \gamma^2 \sum_{k=0}^{t-1} r_k^2$. Also, if $\gamma_k^2 \leq B > 0$ for all $k$ then $\mathrm{Var}[V_t] = O(t)$. Next, assume that $V_0 = 0$ and consider the standardized value, defined by

$$Z_t = \frac{V_t - \mathrm{E}[V_t]}{\sqrt{\mathrm{Var}[V_t]}}. \tag{3.7}$$

Using Chebyshev's Inequality, we find that

$$\begin{aligned}\mathcal{P}[V_t \le 0] &= \mathcal{P}\left[Z_t \le \frac{-\mathrm{E}[V_t]}{\sqrt{\mathrm{Var}[V_t]}}\right] \\ &\le \mathcal{P}\left[|Z_t| \ge \frac{\mathrm{E}[V_t]}{\sqrt{\mathrm{Var}[V_t]}}\right] \\ &\le \frac{\mathrm{Var}[V_t]}{(\mathrm{E}[V_t])^2} = \frac{\sum_{k=0}^{t-1} r_k^2 \gamma_k^2}{\left[\sum_{k=0}^{t-1} r_k\right]^2} \longrightarrow 0,\end{aligned}$$

as $t \to \infty$, whenever $\mathrm{E}[V_t] = O(t)$ (which is true when $r_k$ is bounded for all $k$) and $\mathrm{Var}[V_t] = O(t^{2-\delta})$ for some $\delta > 0$. Thus, we have shown that there exists an eventual arbitrage, defined as follows.

**Definition 3.** We will say that a strategy $\boldsymbol{\phi}$ admits an *eventual arbitrage* whenever $V_0 = 0$ and

$$\lim_{t\to\infty} \mathcal{P}[V_t > 0] = 1. \tag{3.8}$$

It is instructive to state the conditions on the model that lead to this result. First, the market must be self-financing. Second, the inverse $\boldsymbol{\Sigma}_t^{-1}$ must exist and $\boldsymbol{\mu}_t{}'\boldsymbol{\Sigma}_t^{-1}\boldsymbol{\mu}_t > 0$. Third, $\mathrm{Var}[V_t] = O(t^{2-\delta})$ for some $\delta > 0$ which means that the variance is growing slowly relative to the expectation of $V_t$.

**Example 1.** Suppose that $P(t) = P(0)\exp\{(\mu - {}^1\!/\!_2\sigma^2)t + \sigma W(t)\}$ where $W(t)$ is a Weiner process and $\mathcal{H}_t = \sigma\{P(s) : 0 \le s \le t\}$. In this case, the geometric Weiner proces has $\mu_t = \mathrm{E}[\Delta_t|\mathcal{H}_t] = P(t)(e^\mu - 1)$ and $\phi_t = c[P(t)(e^\mu - 1)]^{-1}$, as long as $\mu \neq 0$. Moreover $\Sigma_t = \mathrm{Var}[\Delta_t|\mathcal{H}_t] = [P(t)]^2 e^{2\mu+\sigma^2}$. Thus $\mathrm{Var}[\phi_t \cdot \Delta_t|\mathcal{H}_t] = c^2 e^{\sigma^2}$ and $\mathrm{Var}[V_t] = tc^2 e^{\sigma^2}$. Moreover, under the optimal strategy we find that $V_t$ is a sum of independent and identically distributed random variables and so the Central Limit Theorem implies that $Z_t$ converges in distribution to a standard normal variate and $V_t$ admits an eventual arbitrage.

**Example 2.** Using a day trader's cost model with $C_t > 0$ and $D_t = 0$ for all $t$, we find that $V_t$ will yield an eventual arbitrage whenever the trading strategy is *exploitable*. That is,

$$\mathrm{E}\left[\boldsymbol{\phi}_t \cdot \boldsymbol{\Delta}_t - |\boldsymbol{\phi}_t| \cdot \boldsymbol{c}_t \mid \mathcal{H}_t\right] > 0, \quad \forall\, t.$$

## 4 The Yield Data

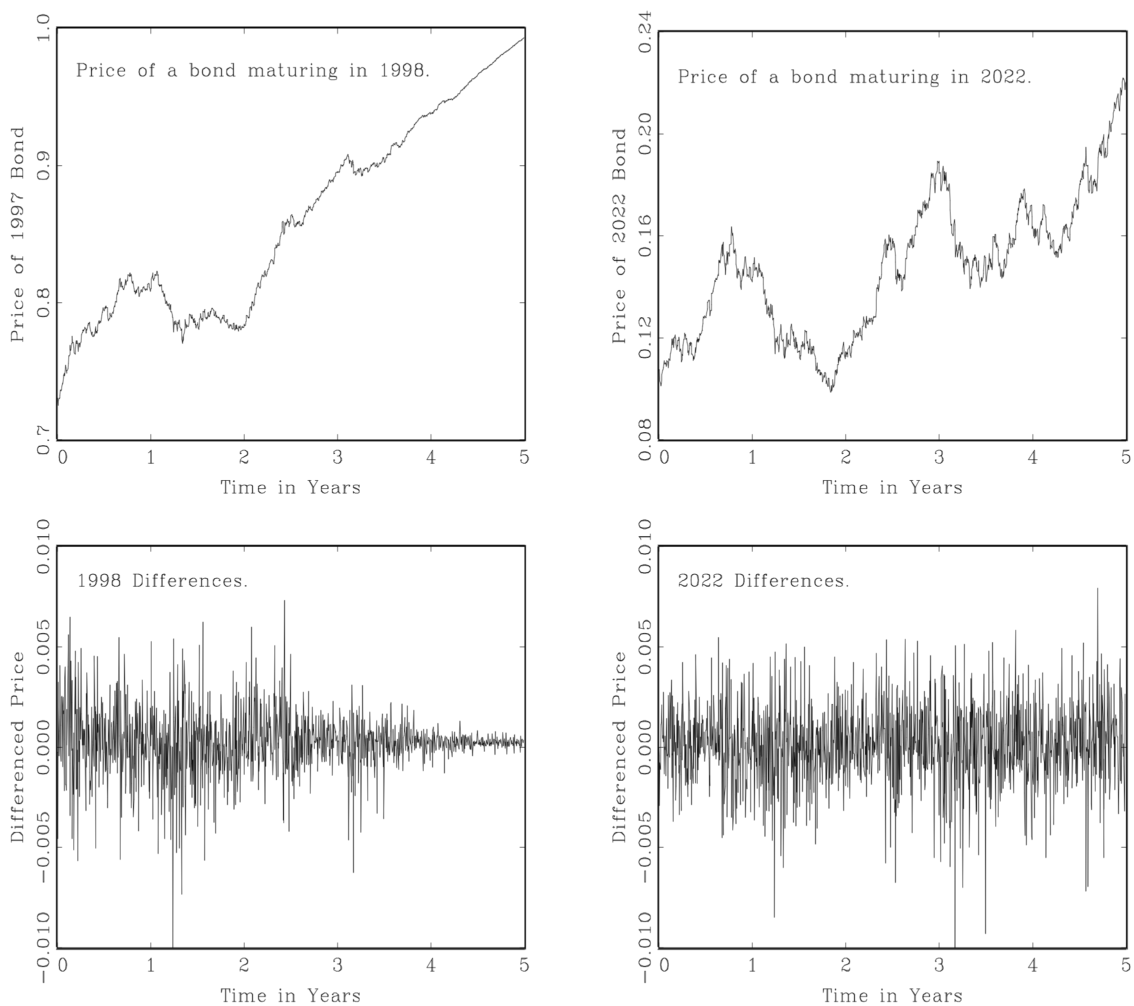
In this section, we present the yield data that we used to test our hypothesis. One of the best sources of data for yield rates are the yield rates on *U.S. Treasury Strips*, as reported by *Bear, Stearns & Co.* via *Software Technologies Inc.* and published in the *Wall Street Journal*. Strips are pure discount bonds that arise from stripping the principal and coupons from government bonds and trading these strips separately. This data was used previously Carrière (2001).

Let us describe the data in detail. In this article we used yield rates from all the trading days in the years of 1993 to 1997, inclusive. In all, we had 1250 trading days with 250 trading days per year. For each trading day, the data consisted of the bid and asked yield rates for stripped coupon interest, stripped Treasury Bond principal, and stripped Treasury Note principal at various maturities. For our purposes, the yield of the bond at a fixed maturity was the average of the bid and asked yields for all coupon and principal strips with that maturity. In all, we had 100 bonds with distinct maturities. Specifically, these bonds have maturity dates at every three months during the 25 year period of 1998 to 2022, inclusive. Let $t_i$ for $i = 1, 2, \ldots, 1250$, denote the trading times and let $T_j$ for $j = 1, 2, \ldots, 100$ denote the maturity dates. Thus, for each time and maturity we observed the yield rates, $y(t_i, T_j)$. The corresponding prices are calculated as follows:

$$P(t_i, T_j) = \exp\left\{-(T_j - t_i)\, y(t_i, T_j)\right\}. \tag{4.1}$$

To get an idea of the type of data that we have, we present Figure 1 where $P(t_i, T_1)$, $\Delta_{i1} = P(t_{i+1}, T_1) - P(t_i, T_1)$ and $P(t_i, T_{100})$, $\Delta_{i,100} = P(t_{i+1}, T_{100}) - P(t_i, T_{100})$ are plotted versus $t_i$.

Figure 1: The first column shows the prices and their differences versus time for a strip that matures in February, 1998. The second column shows the prices and their differences versus time for a strip that matures in February, 2022.



Note that the price of our bonds approaches one when we approach maturity. Also note that the volatility approaches zero when approaching maturity. All the graphs and calculations in this article were done with *Gauss*, a matrix programming language.

# 5 Calibration of the Model

In this section, we present a model of $\boldsymbol{\mu}_t$ and $\boldsymbol{\Sigma}_t$ that is calibrated with half the data, while the other half is used to construct a variance-minimizing and self-financing strategy. We will report on the portfolio value $V_t$, as it evolves over our prediction period. For each $k = 1, 2, \ldots, p = 100$ we assume that

$$\Delta_{ik} = \boldsymbol{a}_k' \boldsymbol{\Delta}_{i-1} + d(t_i, T_k)\sigma_k \epsilon_{ik}, \tag{5.1}$$

where $d(t, T)$ is a function (possibly stochastic) having the property, $d(t, t) = 0$, and where $\sigma_k \in \mathbb{R}$ is fixed and non-stochastic. We also require that $\mathrm{E}[\epsilon_{ik}|\mathcal{H}_{t_i}] = 0$, and $\mathrm{Var}[\epsilon_{ik}|\mathcal{H}_{t_i}] = 1$, and

$$\mathrm{Cov}[\epsilon_{ik}, \epsilon_{il}|\mathcal{H}_{t_i}] = \rho_{kl} \in [-1, 1].$$

Some examples of $d(t, T)$ are

$$\begin{aligned} d(t,T) &= (T-t)^{\gamma}, \\ d(t,T) &= (-\ln[P(t,T)])^{\kappa}, \\ d(t,T) &= [P(t,T)]^{\eta}(T-t)^{\gamma}, \\ d(t,T) &= ([P(t,T)]^{\eta}(-\ln[P(t,T)])^{\kappa}, \end{aligned} \tag{5.2}$$

where $\gamma > 0$, $\eta > 0$ and $\kappa > 0$. This is a classical regression problem where a generalized least-squares solution is optimal. We denote the estimates as follows: $\widehat{\boldsymbol{a}}_k$, $\widehat{\sigma}_k$, $\widehat{\rho}_{kl}$. The estimators are:

$$\begin{aligned} \widehat{\boldsymbol{a}}_k &= [\boldsymbol{X}'\boldsymbol{X}]^{-1}\boldsymbol{X}'\boldsymbol{Y}_k, \\ \widehat{\sigma}_k^2 &= \frac{1}{N}\boldsymbol{Y}_k'\left[\mathrm{I} - \boldsymbol{X}[\boldsymbol{X}'\boldsymbol{X}]^{-1}\boldsymbol{X}'\right]\boldsymbol{Y}_k, \\ \widehat{\rho}_{kl} &= \frac{\boldsymbol{Y}_k'[\mathrm{I} - \boldsymbol{X}[\boldsymbol{X}'\boldsymbol{X}]^{-1}\boldsymbol{X}']\boldsymbol{Y}_l}{N\,\widehat{\sigma}_k\,\widehat{\sigma}_l}, \end{aligned} \tag{5.3}$$

where $N$ is the number of observations and

$$\boldsymbol{Y}_k = \left[\frac{\Delta_{1k}}{d(t_1, T_k)}, \ldots, \frac{\Delta_{Nk}}{d(t_N, T_k)}\right]'. \tag{5.4}$$

To define the design matrix $\boldsymbol{X}$, we first define

$$\boldsymbol{x}_k = \left[\frac{\Delta_{0,k}}{d(t_0, T_k)}, \ldots, \frac{\Delta_{N-1,k}}{d(t_{N-1}, T_k)}\right]'. \tag{5.5}$$

Thus

$$\boldsymbol{X} = [\ \boldsymbol{x}_1, \ldots, \boldsymbol{x}_p\ ]. \tag{5.6}$$

Let $\mathrm{A} = [\boldsymbol{a}_1, \ldots, \boldsymbol{a}_p]$. In this model,

$$\boldsymbol{\mu}_t = \mathrm{E}[\boldsymbol{\Delta}_t | \mathcal{H}_t] = \mathrm{A}'\boldsymbol{\Delta}_{t-1} \tag{5.7}$$

and $\boldsymbol{\Sigma}_t = \text{Var}[\boldsymbol{\Delta}_t|\mathcal{H}_t] = \{\text{Cov}[\Delta_{tk}, \Delta_{tl}|\mathcal{H}_t]\}_{k,l=1,\dots,p}$ where $\text{Cov}[\Delta_{tk}, \Delta_{tl}|\mathcal{H}_t] = d(t,T_k)d(t,T_l)\sigma_k\sigma_l\rho_{kl}$. Therefore,

$$\begin{aligned} \widehat{\boldsymbol{\mu}}_t &= \widehat{\mathrm{A}}'\boldsymbol{\Delta}_{t-1}, \\ \widehat{\boldsymbol{\Sigma}}_t &= \{\, d(t,T_k)\, d(t,T_l)\, \widehat{\sigma}_k\, \widehat{\sigma}_l\, \widehat{\rho}_{kl}\,\}_{k,l=1,\dots,p}\,, \\ \widehat{\boldsymbol{\phi}}_t &= \frac{r_t\, \widehat{\boldsymbol{\Sigma}}_t^{-1} \widehat{\boldsymbol{\mu}}_t}{\widehat{\boldsymbol{\mu}}_t' \widehat{\boldsymbol{\Sigma}}_t^{-1} \widehat{\boldsymbol{\mu}}_t}. \end{aligned} \tag{5.8}$$

Note that

$$\widehat{\sigma}_k\, \widehat{\sigma}_l\, \widehat{\rho}_{kl} = \frac{1}{N}\boldsymbol{Y}_k' \left[\mathrm{I} - \boldsymbol{X}[\boldsymbol{X}'\boldsymbol{X}]^{-1}\boldsymbol{X}'\right] \boldsymbol{Y}_l \tag{5.9}$$

*Prediction*

Using the estimates $\widehat{\boldsymbol{\phi}}_t$, which are calibrated with the data at time $t = 1, 2, \dots N$, we predict according to the optimal rule for the next $N$ observations and observe the value of our self-financing portfolio with a starting value of $V_N = 0$. Thus,

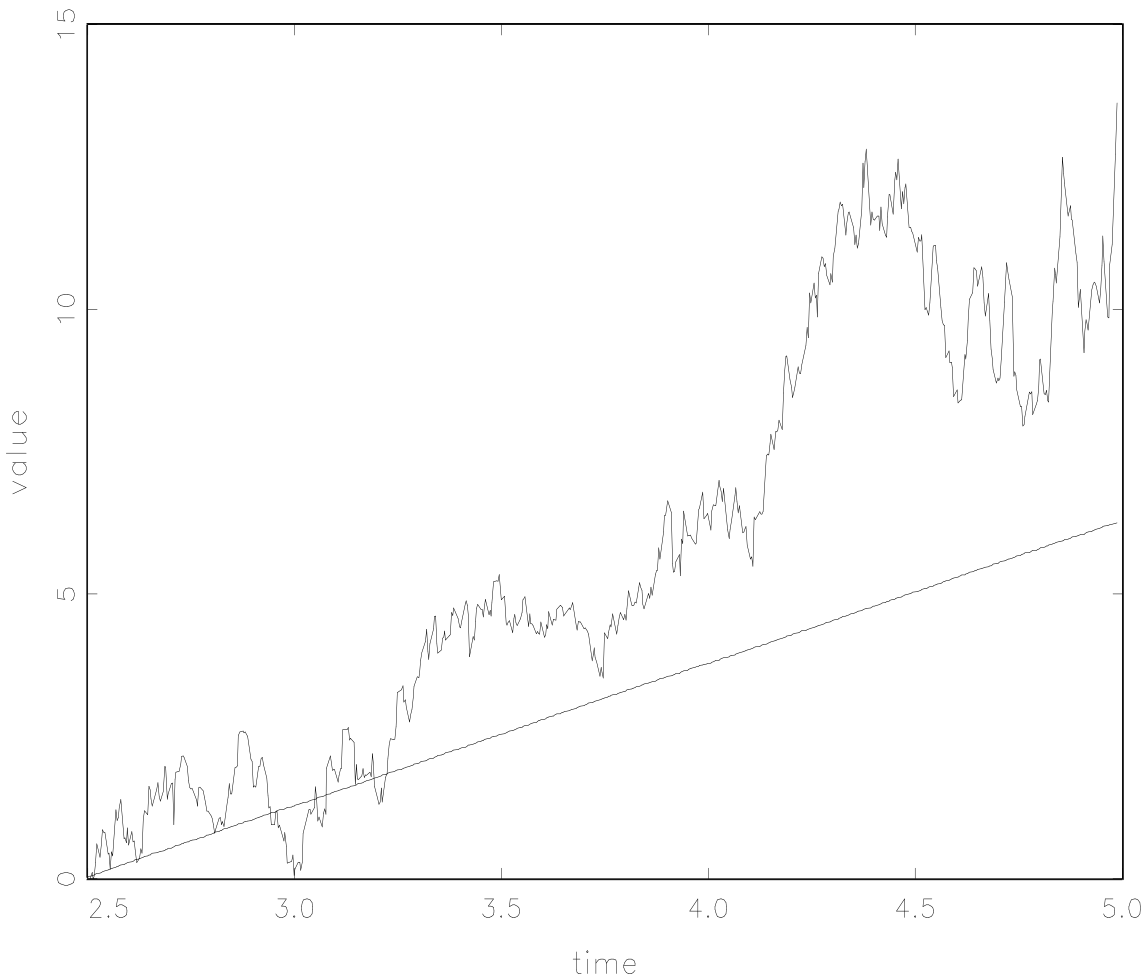
$$V_{2N} = \sum_{k=N}^{2N} \widehat{\boldsymbol{\phi}}_k \cdot \boldsymbol{\Delta}_k. \tag{5.10}$$

The result is plotted in Figure 2, where we find a succesful result, That is the fund value is increasing linearly since $r_k = r$ for all $k$. In this demonstration, we assumed that $d(t,T) = -P(t,T)\ln P(t,T)$. Finally. an estimate of the average return per unit traded is

$$\frac{\sum_{k=N}^{2N} \widehat{\boldsymbol{\phi}}_k \cdot \boldsymbol{\Delta}_k}{\sum_{k=N}^{2N} |\widehat{\boldsymbol{\phi}}_k| \cdot \mathbf{1}} = 2.72 \times 10^{-6}, \tag{5.11}$$

since $|\widehat{\boldsymbol{\phi}}_k| \cdot \mathbf{1}$ is the total number of trades at time $k$. With 250 trading days in a year, we can calculate the nominal annual rate of return when trading with a variance-minimizing strategy. The rate is

Figure 2: Pot of the Value from a Self-Financing Portfolio.



$100 \times 250 \times 2.72/10^6 = .068\%$, which is miniscule. Thus, it seems very unlikely that this market is exploitable. Thus, a lower bound on the cost of buying or selling a million dollars in bonds is $\$2.72$. This bound can be improved by developping better predictive models.

## References

Carrière J. (2001) "A Gaussian Process of Yield Rates Calibrated with Strips." *North American Actuarial Journal*, Vol. 5, No. 3. pp 19-30.

Musiela, M. and Rutkowski, M. (1998): *Martingale Methods in Financial Modelling*. New York, N.Y.: Springer-Verlag.

Panjer, H. (Editor), Boyle, P., Cox, S., Dufresne, D., Gerber, H., Mueller, H., Pedersen, H., Pliska, S., Sherris, M., Shiu, E., Tan, K.S. (1998): *Financial Economics: With Applications to Investments, Insurance and Pensions*. Schaumburg, Ill.: The Actuarial Foundation.

*Last revised: August 20, 2002*