Algorithmic Trading of Co-integrated Assets

Álvaro Cartea[a], Sebastian Jaimungal[b]

[a] *Department of Mathematics, University of Oxford, Oxford, UK*
Oxford-Man Institute of Quantitative Finance, Oxford, UK

[b] *Department of Statistical Sciences, University of Toronto, Toronto, Canada*

Abstract

We assume that the drift in the returns of asset prices consists of an idiosyncratic component and a common component given by a co-integration factor. We analyze the optimal investment strategy for an agent who maximizes expected utility of wealth by dynamically trading in these assets. The optimal solution is constructed explicitly in closed-form and is shown to be affine in the co-integration factor. We calibrate the model to three assets traded on the Nasdaq exchange (Google, Facebook, and Amazon) and employ simulations to showcase the strategy's performance.

Keywords: Pairs trading, algorithmic trading, high-frequency trading, co-integration, short-term alpha, stochastic control.

1. Introduction

The success of many trading algorithms depends on the quality of the predictions of stock price movements. Predictions of the price of a single stock are generally less accurate than predictions of a portfolio of stocks. A classical strategy which makes the most of the predictability of the joint, rather than the individual, behavior of two assets is 'pairs trading' where a portfolio consisting of a linear combination of two assets is traded. At the heart of the strategy is how the two assets co-move. As an example, take two assets whose spread, that is the difference between their prices, exhibits a marked pattern and deviations from it are temporary. Pairs trading algorithms profit from betting on the empirical fact that spread deviations tend to return to their historical or predictable level.

In this paper we derive the optimal trading strategy for an agent who takes positions in n

Email addresses: a.cartea@ucl.ac.uk (Álvaro Cartea), sebastian.jaimungal@utoronto.ca (Sebastian Jaimungal)

co-integrated assets. At the core of the strategy is to profit from the structural dependence in the assets' price dynamics. We assume that the drifts of the assets are co-integrated and develop an algorithmic trading strategy where the investor maximizes expected utility of wealth. We provide an explicit closed-form expression for the optimal (dynamic) investment strategy and show that it is affine in the co-integration factor. Furthermore, we use trading (ITCH) data from the Nasdaq exchange to calibrate the model and then use simulations to illustrate how the strategy performs when the investor takes positions in three assets: Google, Facebook, and Amazon.

There is a large number of papers that study pairs trading strategies and their contribution is either on the statistical aspects and modelling of the relationship between pairs of assets, whilst few others look at how to dynamically take positions in the pairs. For example, Elliott et al. (2005) propose a mean-reverting Gaussian Markov chain model for the spread of a pair of assets. Duan and Pliska (2004) propose a model for co-integrated asset prices and focus on the valuation of options.

One of the early papers which employs stochastic control techniques to trade pairs of co-integrated assets is that of Mudchanatongsuk et al. (2008). The authors model the log-relationship between a pair of stock prices as an Ornstein-Uhlenbeck process and use this to formulate a portfolio stochastic control problem. More recently, Leung and Li (2015) study the optimal timing strategies for trading a mean-reverting price spread. The authors formulate an optimal double stopping problem to analyze the timing to start and subsequently liquidate the position subject to transaction costs. Lei and Xu (2015) analyze a multiple entry-exit problem on a pair of co-integrated assets. The authors recast the sequence of optimal stopping problems as variational inequalities and perform extensive numerical simulations (as well as calibrate to a pair of dual-listed Chinese stocks) to illustrate how the optimal strategy behaves. The work of Ngo and Pham (2016) considers two correlated assets whose spread is modelled by a mean-reverting process with stochastic volatility and show how the investor switches between holding no stocks, long one short the other stock, and vice-versa.

Our paper is closest to that of Tourin and Yan (2013) who develop an optimal portfolio strategy to invest in two risky assets and the money market account. Tourin and Yan assume that log-prices are co-integrated and find, in closed-form, the dynamic trading strategy that maximizes the investor's expected utility of wealth. In our model we generalize Tourin and Yan to allow the investor to trade in m co-integrated assets and provide an explicit closed-form solution of the dynamic trading strategy. We assume that the drift of asset returns consists of an idiosyncratic and a common drift component. The common component, which we label short-term alpha, is a zero-mean reverting process which is an essential source of profits in the trading strategy – it determines how to benefit from short-lived investment opportunities in the collection of assets.

The remainder of this paper is organized as follows. In Section 2 we present the stock price dynamics of the co-integrated assets. Section 3 develops the investor's dynamic optimization problem and derives in closed-form the optimal investment strategy. In Section 4 we employ high-frequency data from three stocks traded in Nasdaq to show the performance of the strategy. Section 5 concludes and we collect proofs in the Appendix.

2. Co-integrated log prices with short-term Alpha

We assume that the drift of asset returns consists of an idiosyncratic and a common component. The firm specific component is a result of factors that only affect the individual firm, and the common component is the result of industry or sector specific factors which affect a collection of assets. Thus, suppose that we have a collection of risky assets whose vector of prices $\boldsymbol{Y} = (Y_t^1, \ldots, Y_t^n)_{0\le t\le T}$ satisfy the coupled system of SDEs

$$\frac{dY_t^k}{Y_t^k} = (\nu_k + \delta_k\, \alpha_t\,)\, dt + \sum_{i=1}^{n} \sigma_{ki}\, dW_t^i\,, \tag{1}$$

where the idiosyncratic drift component ν_k is a constant, and the common drift component is

$$\alpha_t = \sum_{i=1}^{n} a_i\, \log Y_t^i\,,$$

where a_i are constants, and δ_k is a firm specific constant that scales how the common drift component affects each individual asset.

Furthermore, W_t^i are standard Brownian motions independent of each other, arranged in a vector $\boldsymbol{W} = (W_t^1, \ldots, W_t^n)_{0\le t\le T}$, and σ_{ki} are non-negative constants. One can see that the instantaneous covariance, loosely interpreted as

$$\mathbb{C}\left[\frac{dY_t^i}{Y_t^i}, \frac{dY_t^j}{Y_t^j}\,\middle|\, \mathcal{F}_t\right]\,, \tag{2}$$

between assets i and j is given by

$$\Omega_{ij} = \sum_{k=1}^{n} \sigma_{ik}\sigma_{kj}\,. \tag{3}$$

Thus, we denote by $\boldsymbol{\sigma}$ the matrix whose elements are σ_{ij} which is the Cholesky decomposition of the instantaneous variance-covariance matrix $\boldsymbol{\Omega}$ so that $\boldsymbol{\Omega} = \boldsymbol{\sigma}\boldsymbol{\sigma}'$, where $'$ denotes the transpose operator. We assume that there are no redundant degrees of freedom, so that $\boldsymbol{\sigma}$ is invertible.

Now we show that α_t acts as a co-integration factor. When $\alpha_t = 0$ all assets are correlated geometric Brownian motions with drift ν_k. In general, however, in equity markets, α_t will be non-zero and deviations from its long-term or predictable level quickly subside. This behavior in the drift component is also referred to as short-term alpha because it represents short-term deviations from the assets' expected return, see Cartea et al. (2014).

We justify calling α_t a co-integration factor by demonstrating that it is indeed a mean-reverting process. First, note that the log-prices satisfy the SDEs

$$d\log Y_t^k = \left(\nu_k + \delta_k\,\alpha_t - \tfrac{1}{2}\Omega_{kk}\right)\,dt + \sum_{i=1}^{n} \sigma_{ki}\,dW_t^i\,, \tag{4}$$

which we use to compute the differential of α_t:

$$d\alpha_t = \sum_{k=1}^{n} a_k\left(\nu_k + \delta_k\,\alpha_t - \tfrac{1}{2}\Omega_{kk}\right)\,dt + \sum_{k=1}^{n} a_k \sum_{i=1}^{n} \sigma_{ki}\,dW_t^i\,.$$

Therefore, we can write the short-term alpha component as the mean-reverting process

$$d\alpha_t = \kappa\,(\theta - \alpha_t)\,dt + \boldsymbol{a}'\boldsymbol{\sigma}\,d\boldsymbol{W}_t\,, \tag{5}$$

where

$$\boldsymbol{a} = (a_1, \ldots, a_n)'\,, \tag{6}$$

$$\kappa = -\sum_{k=1}^{n} a_k\,\delta_k = -\boldsymbol{\delta}'\,\boldsymbol{a}\,, \qquad \text{and} \tag{7}$$

$$\theta = \frac{\sum_{k=1}^{n} a_k \Omega_{kk}}{2\sum_{k=1}^{n} a_k\,\delta_k} - \frac{\sum_{k=1}^{n} a_k\,\nu_k}{\sum_{k=1}^{n} a_k\,\delta_k} = \tfrac{1}{2}\frac{\mathrm{Tr}(\boldsymbol{A}\boldsymbol{\Omega})}{\boldsymbol{\delta}'\,\boldsymbol{a}} - \frac{\boldsymbol{\nu}'\,\boldsymbol{a}}{\boldsymbol{\delta}'\,\boldsymbol{a}}\,. \tag{8}$$

Here, the operator $\mathrm{Tr}(\cdot)$ denotes the trace of the matrix in braces and $\boldsymbol{A} = \mathrm{diag}(\boldsymbol{a})$ is a diagonal matrix whose diagonal entries are $\boldsymbol{a}$. To ensure that the model does indeed describe a mean-reverting process, as opposed to a mean-repelling one, we assume that $\boldsymbol{\delta}'\,\boldsymbol{a} < 0$ – something that is also borne out by the results when the model is calibrated to Nasdaq data.

By looking at the dynamics (5), we can see that α mean-reverts at a rate which depends on the various strengths of the impact that α has on each asset (through $\boldsymbol{\delta}$) as well as the strength of the log-asset price's contribution to α itself (through $\boldsymbol{a}$). The mean-reversion level θ depends on the ratio of the volatility relative to the impact each component has on the drift of the assets and on the idiosyncratic drift of the assets.

Another alternative representation of the model stems from inserting the expression for α_t directly into the SDE for $\log Y_t$. In this case, we have

$$d\log Y_t^k = \left(\nu_k - \tfrac{1}{2}\Omega_{kk} + \delta_k \textstyle\sum_{i=1}^{n} a_i\,\log Y_t^i\right)\,dt + \textstyle\sum_{i=1}^{n} \sigma_{ki}\,dW_t^i\,, \tag{9}$$

so that if we let $\boldsymbol{Z}_t = \log \boldsymbol{Y}_t$, where the log is interpreted componentwise, then

$$d\boldsymbol{Z}_t = (\boldsymbol{c} - \boldsymbol{B}\,\boldsymbol{Z}_t)\;dt + \boldsymbol{\sigma}\,d\boldsymbol{W}_t\,, \tag{10}$$

with

$$\boldsymbol{B} = -\begin{pmatrix} \delta_1\,a_1 & \delta_1\,a_2 & \dots & \delta_1\,a_n \\ \vdots & \vdots & & \vdots \\ \delta_n\,a_1 & \delta_n\,a_2 & \dots & \delta_n\,a_n \end{pmatrix}, \quad \text{and} \quad \boldsymbol{c} = \begin{pmatrix} \nu_1 - \frac{1}{2}\,\Omega_{11} \\ \vdots \\ \nu_n - \frac{1}{2}\,\Omega_{nn} \end{pmatrix}. \tag{11}$$

In this representation, we see that the log-prices are an order 1 vector-autoregressive model (VAR(1)). Although this representation is in some sense more compact, we find in the next section that the short-term alpha version of the model is a preferable representation for solving the agent's optimal investment problem. Note that here, $\boldsymbol{B}$ is singular and contains exactly one positive eigenvalue due to the singular co-integration factor that we incorporate into the model. If we have $m \leq n$ co-integration factors then there will be in general m positive eigenvalues.

3. Agent's investment optimization problem

The agent's goal is to take positions in the assets to maximize the expected utility of wealth and these positions are constantly revised as time evolves to ensure that the strategy is dynamically optimal. We let $\pi = (\pi_t^0, \pi_t^1, \dots, \pi_t^n)_{0\leq t\leq T}$ denote the dollar value invested in the riskless (π_t^0) and risky assets $(\pi_t^1, \dots, \pi_t^n)$, and let $X^{\boldsymbol{\pi}} = (X_t^{\boldsymbol{\pi}})_{0\leq t\leq T}$ denote the agent's (controlled) wealth process. With this convention, the number of units m_t^k the agent holds in asset k is $m_t^k = \pi_t^k / Y_t^k$. Hence, the wealth process can be written as

$$X_t^{\boldsymbol{\pi}} = \sum_{k=0}^{n} \pi_t^k = \sum_{k=0}^{n} m_t^k\,Y_t^k\,, \tag{12}$$

so that

$$\begin{aligned} dX_t^{\boldsymbol{\pi}} &= \sum_{k=0}^{n} d(m_t^k\,Y_t^k) = \sum_{k=0}^{n} m_t^k\,dY_t^k = \sum_{k=1}^{n} \pi_t^k\,\frac{dY_t^k}{Y_t^k} & (13)\\ &= \sum_{k=1}^{n} \pi_t^k \left\{ (\nu_k + \delta_k\,\alpha_t\,)\,dt + \sum_{i=1}^{n} \sigma_{ki}\,dW_t^i \right\} = \boldsymbol{\pi}_t'\,(\boldsymbol{\nu} + \boldsymbol{\delta}\,\alpha_t)\;dt + \boldsymbol{\pi}_t'\boldsymbol{\sigma}d\boldsymbol{W}_t\,, & (14) \end{aligned}$$

where the second equality in (13) follows from the usual self-financing constraint – which can be interpreted as the change in the wealth process due to the change in each asset's value, assuming the positions are held fixed over a small interval of time. The third equality is obtained by assuming that the time horizon is short enough that interest rates are zero (so that $dY_t^0 = 0$), but long enough that the geometric model we employ is required. The last equality in (14) is a rewrite of the equations using vector/matrix notation.

Next, the agent optimizes her dollar value in assets directly, rather than the rate of trading (as in many formulation of algorithmic trading problems involving liquidation, see, e.g., Cartea et al. (2015)), and has exponential utility $u(x) = -e^{-\gamma x}$. Hence, her performance criteria is given by

$$H^{\boldsymbol{\pi}}(t,x,\boldsymbol{y}) = \mathbb{E}_{t,x,\boldsymbol{y}}\left[-\exp\left(-\gamma\, X_T^{\boldsymbol{\pi}}\right)\right] , \tag{15}$$

where $\mathbb{E}_{t,x,\boldsymbol{y}}[\cdot]$ represents expectation conditional on $X_t^{\boldsymbol{\pi}} = x$ and $\boldsymbol{Y}_t = \boldsymbol{y}$, and T is the terminal date at which the agent liquidates her holdings and stops trading. The agent's value function is

$$H(t,x,\boldsymbol{y}) = \sup_{\boldsymbol{\pi}\in\mathcal{A}} H^{\boldsymbol{\pi}}(t,x,\boldsymbol{y})\,, \tag{16}$$

where the set of admissible strategies $\mathcal{A}$ are those for which

$$\mathbb{E}\left[\sum_{k=0}^{n}\int_0^T (\pi_u^k)^2\,du\right] < \infty\,. \tag{17}$$

Alternatively, we can enforce the condition that each component of $\boldsymbol{\pi}$ is $\mathbb{P}$-a.s. bounded.

Applying the dynamic programming principle leads to the dynamic programming equation (DPE) which the value function should satisfy:

$$\begin{aligned}&\partial_t H + (\boldsymbol{\nu}' + \alpha\,\boldsymbol{\delta}')\,\boldsymbol{\mathcal{D}}_y H + \tfrac{1}{2}\boldsymbol{\mathcal{D}}_{yy}^{\boldsymbol{\Omega}} H \\ &\qquad + \sup_{\boldsymbol{\pi}}\left\{\boldsymbol{\pi}'\,(\boldsymbol{\nu} + \boldsymbol{\delta}\,\alpha)\;\partial_x H + \tfrac{1}{2}(\boldsymbol{\pi}_t'\boldsymbol{\Omega}\boldsymbol{\pi})\,\partial_{xx}H + \boldsymbol{\pi}'\Omega\,\boldsymbol{\mathcal{D}}_{xy}H\right\} = 0\,,\end{aligned} \tag{18}$$

subject to the terminal condition $H(T,x,\boldsymbol{y}) = -e^{-\gamma\,x}$, and where $\alpha = \boldsymbol{a}'\log\boldsymbol{y}$ (the log being interpreted componentwise) represents the state of the co-integration process, and the following linear differential operators were introduced:

$$\boldsymbol{\mathcal{D}}_y H = (y^1\partial_{y^1}H, \dots, y^n\partial_{y^n}H)'\,, \tag{19}$$

$$\boldsymbol{\mathcal{D}}_{yy}^{\boldsymbol{\Omega}} H = \textstyle\sum_{i,j=1}^n y^i\Omega_{ij}y^j\partial_{y^iy^j}H\,, \tag{20}$$

$$\boldsymbol{\mathcal{D}}_{xy} H = (y^1\partial_{xy^1}H, \dots, y^n\partial_{xy^n}H)'\,. \tag{21}$$

Proposition 1. *Let the investor's value function satisfy equation* (18). *Then the optimal investment strategy in feedback form is*

$$\boldsymbol{\pi}^* = -\frac{\boldsymbol{\Omega}^{-1}\,(\boldsymbol{\nu} + \boldsymbol{\delta}\,\alpha)\;\partial_x H + \boldsymbol{\mathcal{D}}_{xy}H}{\partial_{xx}H}\,, \tag{22}$$

and the DPE reduces to the non-linear partial differential equation

$$\partial_t H + (\boldsymbol{\nu} + \alpha\,\boldsymbol{\delta})'\,\boldsymbol{\mathcal{D}}_y H + \tfrac{1}{2}\boldsymbol{\mathcal{D}}_{yy}^{\boldsymbol{\Omega}} H - \frac{\boldsymbol{\mathcal{L}}'H\,\boldsymbol{\Omega}^{-1}\boldsymbol{\mathcal{L}}H}{2\,\partial_{xx}H} = 0\,, \tag{23}$$

where the second-order differential operator $\boldsymbol{\mathcal{L}}$ *acts as follows*

$$\boldsymbol{\mathcal{L}}H = (\boldsymbol{\nu} + \boldsymbol{\delta}\,\alpha\,)\,\partial_x H + \boldsymbol{\Omega}\,\boldsymbol{\mathcal{D}}_{xy}H\,. \tag{24}$$

PROOF. Let us focus on the supremum term of the DPE (18) and perform a matrix completion of the square. Specifically, assuming that $\partial_{xx} H \neq 0$, we write

$$
\begin{aligned}
\mathcal{M} &= \boldsymbol{\pi}' \left(\boldsymbol{\nu} + \boldsymbol{\delta}\,\alpha\right)\, \partial_x H + \left(\boldsymbol{\pi}'\boldsymbol{\Omega}\boldsymbol{\pi}\right) \partial_{xx} H + \boldsymbol{\pi}'\boldsymbol{\Omega}\,\boldsymbol{\mathcal{D}}_{xy} H \\
&= \tfrac{1}{2}\partial_{xx} H \left\{ \boldsymbol{\pi}'\boldsymbol{\Omega}\boldsymbol{\pi} + 2\,\boldsymbol{\pi}' \frac{\left(\boldsymbol{\nu} + \boldsymbol{\delta}\,\alpha\right) \partial_x H + \boldsymbol{\Omega}\,\boldsymbol{\mathcal{D}}_{xy} H}{\partial_{xx} H} \right\} \\
&= \tfrac{1}{2}\partial_{xx} H \left\{ \left(\boldsymbol{\pi} + \boldsymbol{\Omega}^{-1}\,\boldsymbol{\mathcal{L}} H\right)' \boldsymbol{\Omega} \left(\boldsymbol{\pi} + \boldsymbol{\Omega}^{-1}\,\boldsymbol{\mathcal{L}} H\right) - \frac{\boldsymbol{\mathcal{L}}' H\,\boldsymbol{\Omega}^{-1}\boldsymbol{\mathcal{L}} H}{(\partial_{xx} H)^2} \right\}, \qquad (25)
\end{aligned}
$$

where the vector-valued linear operator $\boldsymbol{\mathcal{L}}$ acts on H as follows:

$$
\boldsymbol{\mathcal{L}} H = \left(\boldsymbol{\nu} + \boldsymbol{\delta}\,\alpha\,\right) \partial_x H + \boldsymbol{\Omega}\,\boldsymbol{\mathcal{D}}_{xy} H\,, \qquad (26)
$$

and we have used the fact that $\boldsymbol{\Omega}$ is symmetric (since it is a variance-covariance matrix) so that $(\boldsymbol{\Omega}^{-1})' = (\boldsymbol{\Omega}')^{-1} = \boldsymbol{\Omega}^{-1}$. From the above expression we can immediately identify the optimal investment in feedback control form as

$$
\boldsymbol{\pi}^* = -\frac{\boldsymbol{\Omega}^{-1}\left(\boldsymbol{\nu} + \boldsymbol{\delta}\,\alpha\right)\, \partial_x H + \boldsymbol{\mathcal{D}}_{xy} H}{\partial_{xx} H}\,. \qquad (27)
$$

From the above matrix completion of the square, we also have the maximum term

$$
\mathcal{M}^* = -\tfrac{1}{2}\,\frac{\boldsymbol{\mathcal{L}}' H\,\boldsymbol{\Omega}^{-1}\boldsymbol{\mathcal{L}} H}{\partial_{xx} H}\,, \qquad (28)
$$

so that upon reinsertion into the DPE, we obtain the following non-linear partial differential equation (PDE) for the value function:

$$
\partial_t H + \left(\boldsymbol{\nu} + \alpha\,\boldsymbol{\delta}\right)' \boldsymbol{\mathcal{D}}_y H + \tfrac{1}{2}\boldsymbol{\mathcal{D}}^{\boldsymbol{\Omega}}_{yy} H - \frac{\boldsymbol{\mathcal{L}}' H\,\boldsymbol{\Omega}^{-1}\boldsymbol{\mathcal{L}} H}{2\,\partial_{xx} H} = 0\,. \qquad (29)
$$

3.1. Solving the DPE

In the classical Merton problem, where asset prices are geometric Brownian motions, the value function for exponential utility has the form $-e^{-\gamma(x+h(t))}$, where h is a deterministic function of time. Here, due to the presence of the co-integration factor we expect instead that the value function depends also on a combination of prices of all assets in the co-integration factor, i.e. depends on the short-term alpha component. Thus, we propose as trial solution to write the value function as

$$
H(t, x, \boldsymbol{y}) = -\exp\left\{ -\gamma\left[x + h\left(t, \textstyle\sum_{i=1}^n a_i\, \log y^i\right)\right]\right\}, \qquad (30)
$$

for some function $h(t, \alpha)$, with $\alpha = \sum_{i=1}^n a_i\, \log y^i$, and subject to the terminal condition $h(T, \alpha) = 0$.

Proposition 2. *Using the ansatz* (30) *for the value function, the function h satisfies the linear PDE:*

$$\partial_t h - \tfrac{1}{2}\,\mathrm{Tr}(\boldsymbol{A}\,\boldsymbol{\Omega})\,\partial_\alpha h + \tfrac{1}{2}(\boldsymbol{a}'\,\boldsymbol{\Omega}\,\boldsymbol{a})\,\partial_{\alpha\alpha} h + \tfrac{1}{2\gamma}(\boldsymbol{\delta}\,\alpha + \boldsymbol{\nu})'\,\boldsymbol{\Omega}\,(\boldsymbol{\delta}\,\alpha + \boldsymbol{\nu})\,\alpha^2 = 0\,, \tag{31}$$

subject to $h(T, \alpha) = 0$.

Moreover, the optimal investment strategy, in feedback form, reduces

$$\boldsymbol{\pi}^* = \tfrac{1}{\gamma}\,\boldsymbol{\Omega}^{-1}(\boldsymbol{\delta}\,\alpha + \boldsymbol{\nu}) - \boldsymbol{a}\,\partial_\alpha h\,. \tag{32}$$

Proof. First note that differentiation w.r.t y^k acts simply on h, specifically

$$\partial_{y^k} h(t,\alpha) = \frac{a_k}{y^k}\,\partial_\alpha h(t,\alpha)\,, \tag{33}$$

so that

$$\partial_{y^k} e^{-\gamma\,h(t,\alpha)} = -\gamma\,\frac{a_k}{y^k}\,\partial_\alpha h(t,\alpha)\,e^{-\gamma\,h(t,\alpha)}\,, \tag{34}$$

$$\partial_{y^j y^k} h(t,\alpha) = -\gamma\,\partial_{y^j}\left(\frac{a_k}{y^k}\,\partial_\alpha h(t,\alpha)\,e^{-\gamma\,h(t,\alpha)}\right)\,. \tag{35}$$

Hence, for $j \neq k$,

$$\begin{aligned}\partial_{y^j y^k} e^{-\gamma\,h(t,\alpha)} &= -\gamma\,\frac{a_k\,a_j}{y^k\,y^j}\,\partial_\alpha\left(\partial_\alpha h(t,\alpha)\,e^{-\gamma\,h(t,\alpha)}\right)\\ &= -\gamma\,\frac{a_k\,a_j}{y^k\,y^j}\,\left(\partial_{\alpha\alpha} h(t,\alpha) - \gamma\,(\partial_\alpha h(t,\alpha))^2\right)\,e^{-\gamma\,h(t,\alpha)}\,,\end{aligned} \tag{36}$$

while for $j = k$,

$$\begin{aligned}\partial_{y^j y^j} e^{-\gamma\,h(t,\alpha)} &= -\gamma\,\left(-\frac{a_k}{(y^k)^2}\,\partial_\alpha h(t,\alpha)\,e^{-\gamma\,h(t,\alpha)} + \frac{a_k^2}{(y^k)^2}\partial_\alpha\left(\partial_\alpha h(t,\alpha)\,e^{-\gamma\,h(t,\alpha)}\right)\right)\\ &= -\gamma\,\frac{a_k}{(y^k)^2}\left(-\partial_\alpha h(t,\alpha) + a_k\left(\partial_{\alpha\alpha} h(t,\alpha) - \gamma\,(\partial_\alpha h(t,\alpha))^2\right)\right)\,e^{-\gamma\,h(t,\alpha)}\,.\end{aligned} \tag{37}$$

Putting these two results together we can write

$$\begin{aligned}\partial_{y^j y^k} e^{-\gamma\,h(t,\alpha)} =& -\gamma\,\frac{a_k\,a_j}{y^k\,y^j}\,\left(\partial_{\alpha\alpha} h(t,\alpha) - \gamma\,(\partial_\alpha h(t,\alpha))^2\right)\,e^{-\gamma\,h(t,\alpha)}\\ &+ \varsigma_{jk}\,\gamma\,\frac{a_k}{(y^k)^2}\,\partial_\alpha h(t,\alpha)\,e^{-\gamma\,h(t,\alpha)}\,,\end{aligned} \tag{38}$$

where ς_{jk} is the Kronecker delta which equals 1 if $j = k$ and 0 otherwise.

Armed with these results, the various linear differential operators which appear in the non-linear PDE (23) can be written as follows:

$$\boldsymbol{\mathcal{D}}_y H = (-\gamma\, H)\, \boldsymbol{a}\, \partial_\alpha h\,, \tag{39}$$

$$\begin{aligned}\boldsymbol{\mathcal{D}}^{\boldsymbol{\Omega}}_{yy} H &= (-\gamma\, H)\, \textstyle\sum_{i,j=1}^{n} \Omega_{ij}\, a_i\, a_j\, \left(\partial_{\alpha\alpha} h - \gamma(\partial_\alpha h)^2\right) + (\gamma\, H)\, \textstyle\sum_{j=1}^{n} a_j\, \Omega_{jj}\, \partial_\alpha h \\ &= -(\gamma\, H)\, (\boldsymbol{a}'\boldsymbol{\Omega}\boldsymbol{a})\, \big(\partial_{\alpha\alpha} h - \gamma(\partial_\alpha h)^2\big) + (\gamma\, H)\, \mathrm{Tr}(\boldsymbol{A}\,\boldsymbol{\Omega})\, \partial_\alpha h\,. \end{aligned} \tag{40}$$

Furthermore,

$$\boldsymbol{\mathcal{D}}_{xy} H = (\gamma^2\, H)\, \boldsymbol{a}\, \partial_\alpha h\,, \tag{41}$$

$$\boldsymbol{\mathcal{L}} H = (-\gamma\, H)\, (\boldsymbol{\delta}\, \alpha + \boldsymbol{\nu}) + (\gamma^2\, H)\, \boldsymbol{\Omega}\boldsymbol{a}\, \partial_\alpha h\,. \tag{42}$$

Inserting these expressions into the PDE (23), allows us to write

$$\begin{aligned} 0 =\, & (-\gamma H)\partial_t h + (-\gamma H)\, (\boldsymbol{\delta}\, \alpha + \boldsymbol{\nu})'\, \boldsymbol{a}\, \partial_\alpha h \\ & + \tfrac{1}{2}\Big((-\gamma\, H)\, (\boldsymbol{a}'\boldsymbol{\Omega}\boldsymbol{a})\, \big(\partial_{\alpha\alpha} h - \gamma(\partial_\alpha h)^2\big) + (\gamma\, H)\, \mathrm{Tr}(\boldsymbol{A}\,\boldsymbol{\Omega})\, \partial_\alpha h\Big) \\ & - \frac{1}{2\gamma^2\, H}\, \Big((-\gamma\, H)\, (\boldsymbol{\delta}\alpha + \boldsymbol{\nu}) + (\gamma^2\, H)\, \boldsymbol{\Omega}\boldsymbol{a}\, \partial_\alpha h\Big)'\, \boldsymbol{\Omega}^{-1} \Big((-\gamma\, H)\, (\boldsymbol{\delta}\alpha + \boldsymbol{\nu}) + (\gamma^2\, H)\, \boldsymbol{\Omega}\boldsymbol{a}\, \partial_\alpha h\Big)\,. \end{aligned} \tag{43}$$

At this point, there are three important simplifications. First, cancelling $-\gamma\, H$ in all terms, we find that

$$\begin{aligned} 0 =\, & \partial_t h + \ (\boldsymbol{\delta}'\boldsymbol{a})\, \alpha\, \partial_\alpha h \\ & + \tfrac{1}{2}\Big((\boldsymbol{a}'\boldsymbol{\Omega}\boldsymbol{a})\, \big(\partial_{\alpha\alpha} h - \gamma(\partial_\alpha h)^2\big) - \mathrm{Tr}(\boldsymbol{A}\,\boldsymbol{\Omega})\, \partial_\alpha h\Big) \\ & + \frac{1}{2\gamma}\, ((\boldsymbol{\delta}\alpha + \boldsymbol{\nu}) - \gamma\, \boldsymbol{\Omega}\boldsymbol{a}\, \partial_\alpha h)'\, \boldsymbol{\Omega}^{-1}\, ((\boldsymbol{\delta}\alpha + \boldsymbol{\nu}) - \gamma\, \boldsymbol{\Omega}\boldsymbol{a}\, \partial_\alpha h)\,. \end{aligned} \tag{44}$$

Next, expanding the third line, we have

$$\begin{aligned} & ((\boldsymbol{\delta}\alpha + \boldsymbol{\nu}) - \gamma\, \boldsymbol{\Omega}\boldsymbol{a}\, \partial_\alpha h)'\, \boldsymbol{\Omega}^{-1}\, ((\boldsymbol{\delta}\alpha + \boldsymbol{\nu}) - \gamma\, \boldsymbol{\Omega}\, \boldsymbol{a}\, \partial_\alpha h) \\ & \quad = (\boldsymbol{\delta}\, \alpha + \boldsymbol{\nu})'\, \boldsymbol{\Omega}^{-1}\, (\boldsymbol{\delta}\, \alpha + \boldsymbol{\nu})\, \alpha^2 - 2\gamma\, (\boldsymbol{\delta}\, \alpha + \boldsymbol{\nu})'\, \boldsymbol{a}\, \alpha\, \partial_\alpha h + \gamma^2\, \boldsymbol{a}'\, \boldsymbol{\Omega}\boldsymbol{a}\, (\partial_\alpha h)^2\,, \end{aligned} \tag{45}$$

so that upon substituting into the previous expression there are two important cancellations:

1. the non-linear terms containing $(\partial_\alpha h)^2$ from the expansion above and the second line in (44) cancel one another;

2. the terms $\boldsymbol{\delta}'\,\boldsymbol{a}\,\alpha\,\partial_\alpha h$ from the first line in (44) and the above expansion also cancel.

Putting these observations together we find that h satisfies a very simple linear PDE:

$$\partial_t h - \tfrac{1}{2}\,\mathrm{Tr}(\boldsymbol{A}\,\boldsymbol{\Omega})\,\partial_\alpha h + \tfrac{1}{2}(\boldsymbol{a}'\,\boldsymbol{\Omega}\,\boldsymbol{a})\,\partial_{\alpha\alpha} h + \tfrac{1}{2\gamma}(\boldsymbol{\delta}\,\alpha + \boldsymbol{\nu})'\,\boldsymbol{\Omega}\,(\boldsymbol{\delta}\,\alpha + \boldsymbol{\nu})\,\alpha^2 = 0\,, \tag{46}$$

subject to $h(T, \alpha) = 0$.

Returning to the optimal investment (22) in the assets, the ansatz for the value function $H = -\exp\{-\gamma(x + h(t, \alpha))\}$ allows us to write

$$\begin{aligned}\boldsymbol{\pi}^* &= -\,\frac{-\gamma\,H\,\left(\boldsymbol{\Omega}^{-1}(\boldsymbol{\delta}\,\alpha + \boldsymbol{\nu})\right) + (\gamma^2\,H)\,\boldsymbol{a}\,\partial_\alpha h}{\gamma^2\,H} \\ &= \tfrac{1}{\gamma}\,\boldsymbol{\Omega}^{-1}(\boldsymbol{\delta}\,\alpha + \boldsymbol{\nu}) - \boldsymbol{a}\,\partial_\alpha h\,. \end{aligned} \tag{47}$$

Let us recall that the optimal investment in the classical Merton problem (using discounted wealth) is $\frac{1}{\gamma}\boldsymbol{\Omega}^{-1}(\boldsymbol{\nu} - r)$ where r is the risk-free rate and $\boldsymbol{\nu}$ is (in the classical Merton problem) the drift of the GBMs that drives the asset prices. The first term of the above expression is quite similar, since here $r = 0$, and the drift of the assets are $\boldsymbol{\delta}\,\alpha_t + \boldsymbol{\nu}$ – here, however, the drift is stochastic and hence the Merton solution cannot be applied directly. The optimal investment responds to the level of the co-integration factor α_t through the term $\frac{1}{\gamma}\boldsymbol{\Omega}^{-1}\,\boldsymbol{\delta}\,\alpha$.

Moreover, the optimal investment is perturbed from the (naïve modification of the) Merton solution by the second term, i.e. $-\boldsymbol{a}\,\partial_\alpha h$. To understand the exact contribution of this second term we require knowledge of h, so we now turn to finding an explicit solution for h. Remarkably, in subsection 3.1.1 we build an explicit solution of h by looking at a probabilistic interpretation of h, which requires a suitable measure change, and in subsection 3.1.2 we provide the explicit trading strategy. Before proceeding we also note a few features of h. First, from (31), we can see that the solution to h must be quadratic in α. Thus, we also see that the optimal dollar invested in each asset $\boldsymbol{\pi}^*$ will be linear in α. Therefore the optimal strategy depends on the co-integration factor in at most a linear fashion, but the size of the investment will vary with time.

3.1.1. A Probabilistic Representation

The PDE (31) admits an interesting probabilistic interpretation and one which allows us to write in closed-form the agent's investment strategy. First, consider the probability measure change induced by a vector of market prices of risk $\boldsymbol{\lambda}_t$, which induces a new measure $\mathbb{P}^*$ through the Radon-Nikodym derivative

$$\frac{d\mathbb{P}^*}{d\mathbb{P}} = \exp\left\{-\tfrac{1}{2}\int_0^T ||\boldsymbol{\lambda}_s||^2\,ds - \int_0^T \boldsymbol{\lambda}_s'\,d\boldsymbol{W}_s\right\}\,. \tag{48}$$

Girsanov's Theorem implies that the stochastic processes

$$\boldsymbol{W}^*_t = -\int_0^t \boldsymbol{\lambda}_s \, ds + \boldsymbol{W}_t \tag{49}$$

are independent $\mathbb{P}^*$-Brownian motions. Let us choose $\boldsymbol{\lambda}_t$ such that the drift of the traded assets $\boldsymbol{Y}_t$ are martingales. To this end, rewrite the SDE satisfied by the prices of the assets:

$$dY^k_t = (\nu_k + \delta_k \, \alpha_t + \textstyle\sum_{i=1}^n \sigma_{ki} \lambda^i_t) \; Y^k_t \, dt + \sum_{i=1}^n \sigma_{ki} \, Y^k_t \, dW^{*i}_t \, . \tag{50}$$

The martingale condition requires that

$$\sum_{i=1}^n \sigma_{ki} \, \lambda^i_t = -\nu_k - \delta_k \, \alpha_t, \quad k = 1, \dots, n, \quad \Longleftrightarrow \quad \boldsymbol{\sigma} \, \boldsymbol{\lambda}_t = -\boldsymbol{\delta} \, \alpha_t - \boldsymbol{\nu} \, . \tag{51}$$

Therefore, since $\boldsymbol{\sigma}$ is invertible by assumption, we have

$$\boldsymbol{\lambda}_t = -\boldsymbol{\sigma}^{-1} \; (\boldsymbol{\delta} \, \alpha_t + \boldsymbol{\nu}) \; . \tag{52}$$

At this point, we have found the probability measure $\mathbb{P}^*$ which renders the traded assets martingales. Next, we can ask what is the dynamics of the co-integration factor in terms of these new Brownian motions. To this end, from (5) we have

$$d\alpha_t = \; \left(-\tfrac{1}{2} \, \mathrm{Tr}(\boldsymbol{A} \, \boldsymbol{\Omega}) + \boldsymbol{a}'\boldsymbol{\nu} + (\boldsymbol{a}' \, \boldsymbol{\delta})\alpha_t\right) \; dt + \boldsymbol{a}'\boldsymbol{\sigma} \, (\boldsymbol{\lambda}_t \, dt + d\boldsymbol{W}^*_t) \tag{53}$$

so that

$$d\alpha_t = \; -\tfrac{1}{2} \, \mathrm{Tr}(\boldsymbol{A} \, \boldsymbol{\Omega}) \, dt + \boldsymbol{a}'\boldsymbol{\sigma} \, d\boldsymbol{W}^*_t \, . \tag{54}$$

The last equality follows from (52). Surprisingly, although α_t is mean-reverting in the real-world $\mathbb{P}$-measure, it is a Brownian motion under the risk-neutral $\mathbb{P}^*$-measure.

Let us now return to the PDE (31). It can be re-written in the form

$$(\partial_t + \mathcal{L}^{*,\alpha}) \, h + \tfrac{1}{2\gamma} \, (\boldsymbol{\delta} \, \alpha + \boldsymbol{\nu})' \, \boldsymbol{\Omega} \; (\boldsymbol{\delta} \, \alpha + \boldsymbol{\nu}) = 0 \, , \tag{55}$$

subject to $h(T, \alpha) = 0$, where $\mathcal{L}^{*,\alpha}$ is the $\mathbb{P}^*$-infinitesimal generator of α_t and therefore, applying a Feynman-Kac formula we see that $h(t, \alpha)$ can be expressed as the following expectation:

$$h(t, \alpha) = \mathbb{E}^*_{t,\alpha} \left[\tfrac{1}{2\gamma} \int_t^T \; (\boldsymbol{\delta} \, \alpha_s + \boldsymbol{\nu})' \, \boldsymbol{\Omega} \; (\boldsymbol{\delta} \, \alpha_s + \boldsymbol{\nu}) \; ds \right] \, , \tag{56}$$

where $\mathbb{E}^*_{t,\alpha}$ denotes $\mathbb{P}^*$-expectation given that $\alpha_t = \alpha$. Putting this expression for h back into the value function we then have the following relationship:

$$\sup_{\boldsymbol{\pi} \in \mathcal{A}} \mathbb{E}_{t,x,\boldsymbol{y}} \left[-\exp\left(-\gamma \, X^{\boldsymbol{\pi}}_T\right) \right] = -\exp\left\{ -\gamma \, x - \tfrac{1}{2} \, \mathbb{E}^*_{t,x,\boldsymbol{y}} \left[\int_t^T \; (\boldsymbol{\delta} \, \alpha_s + \boldsymbol{\nu})' \, \boldsymbol{\Omega} \; (\boldsymbol{\delta} \, \alpha_s + \boldsymbol{\nu}) \; ds \right] \right\} \, . \tag{57}$$

The future expectation of integrated quadratic form in α_t determines the incremental value of trading on this co-integrated collection of assets. The strength of the contribution increases with volatility (through $\boldsymbol{\Omega}$) as well as through the strength that α_t has on each asset (through the firm specific scaling factors $\boldsymbol{\delta}$).

3.1.2. Explicit Construction of the Optimal Investment Strategy

Based on the representation from the previous section, we can construct an explicit formula for $h(t,\alpha)$. In particular, we have that

$$h(t,\alpha) = \mathbb{E}^*_{t,\alpha}\left[\tfrac{1}{2\gamma}\int_t^T (\boldsymbol{\delta}\,\alpha_s + \boldsymbol{\nu})'\,\boldsymbol{\Omega}\,(\boldsymbol{\delta}\,\alpha_s + \boldsymbol{\nu})\;ds\right] \tag{58}$$

$$\begin{aligned} &= \tfrac{\boldsymbol{\delta}'\boldsymbol{\Omega}\boldsymbol{\delta}}{2\gamma}\,\mathbb{E}^*\left[\int_t^T \left(\alpha - \tfrac{1}{2}\operatorname{Tr}(\boldsymbol{A}\,\boldsymbol{\Omega})(s-t) + \sqrt{\boldsymbol{a}'\boldsymbol{\Omega}\,\boldsymbol{a}}\,(B^*_s - B^*_t)\right)^2 ds\right] \\ &\quad + \tfrac{\boldsymbol{\delta}'\boldsymbol{\Omega}\boldsymbol{\nu}}{\gamma}\int_t^T \left(\alpha - \tfrac{1}{2}\operatorname{Tr}(\boldsymbol{A}\,\boldsymbol{\Omega})(s-t)\right)\,ds + \tfrac{1}{2\gamma}\,(T-t)\,\boldsymbol{\nu}'\,\boldsymbol{\Omega}\,\boldsymbol{\nu}\,, \end{aligned} \tag{59}$$

where we have introduced the process $B^*_t = \frac{1}{\sqrt{\boldsymbol{a}'\boldsymbol{\Omega}\boldsymbol{a}}}\boldsymbol{a}'\boldsymbol{\sigma}\,W^*_t$ which is a standard $\mathbb{P}^*$-Brownian motion. Hence,

$$\begin{aligned} h(t,\alpha) &= \tfrac{\boldsymbol{\delta}'\boldsymbol{\Omega}\boldsymbol{\delta}}{2\gamma}\,\mathbb{E}^*\Bigg[\int_t^T \Bigg\{ \left(\alpha - \tfrac{1}{2}\operatorname{Tr}(\boldsymbol{A}\,\boldsymbol{\Omega})(s-t)\right)^2 \\ &\qquad\qquad + 2\,\left(\alpha - \tfrac{1}{2}\operatorname{Tr}(\boldsymbol{A}\,\boldsymbol{\Omega})(s-t)\right)\sqrt{\boldsymbol{a}'\boldsymbol{\Omega}\boldsymbol{a}}\,(B^*_s - B^*_t) \\ &\qquad\qquad + (\boldsymbol{a}'\boldsymbol{\Omega}\boldsymbol{a})\,(B^*_s - B^*_t)^2\Bigg\}\,ds\Bigg] \\ &\quad + \tfrac{\boldsymbol{\delta}'\boldsymbol{\Omega}\boldsymbol{\nu}}{\gamma}\,\left(\alpha\,\tau - \tfrac{1}{4}\operatorname{Tr}(\boldsymbol{A}\,\boldsymbol{\Omega})\tau^2\right) + \tfrac{1}{2\gamma}\,\tau\,\boldsymbol{\nu}'\,\boldsymbol{\Omega}\,\boldsymbol{\nu}\,, \end{aligned} \tag{60}$$

$$\begin{aligned} &= \tfrac{\boldsymbol{\delta}'\boldsymbol{\Omega}\boldsymbol{\delta}}{2\gamma}\left\{-\tfrac{2}{3\,\operatorname{Tr}(\boldsymbol{A}\,\boldsymbol{\Omega})}\left[\left(\alpha - \tfrac{1}{2}\operatorname{Tr}(\boldsymbol{A}\,\boldsymbol{\Omega})\,\tau\right)^3 - \alpha^3\right] + \tfrac{1}{2}\,(\boldsymbol{a}'\boldsymbol{\Omega}\boldsymbol{a})\,\tau^2\right\} \\ &\quad + \tfrac{\boldsymbol{\delta}'\boldsymbol{\Omega}\boldsymbol{\nu}}{\gamma}\,\left(\alpha\,\tau - \tfrac{1}{4}\operatorname{Tr}(\boldsymbol{A}\,\boldsymbol{\Omega})\tau^2\right) + \tfrac{1}{2\gamma}\,\tau\,\boldsymbol{\nu}'\,\boldsymbol{\Omega}\,\boldsymbol{\nu}\,, \end{aligned} \tag{61}$$

where $\tau = T - t$ and we have used Fubini to interchange the integral and expectation, and used the fact that $\mathbb{E}^*[(B^*_s - B^*_t)^2] = (s-t)$.

In all, we see that $h(t,\alpha)$ is quadratic in α. Consequently, the optimal investment from (32) takes on the explicit form

$$\pi^* = \tfrac{1}{\gamma}\left\{\boldsymbol{\Omega}^{-1}\,(\boldsymbol{\delta}\,\alpha + \boldsymbol{\nu}) + \tfrac{\boldsymbol{\delta}'\boldsymbol{\Omega}\boldsymbol{\delta}}{\operatorname{Tr}(\boldsymbol{A}\,\boldsymbol{\Omega})}\left[\tfrac{1}{2}\operatorname{Tr}(\boldsymbol{A}\,\boldsymbol{\Omega})\,\tau\,\alpha + \tfrac{1}{4}\,(\operatorname{Tr}(\boldsymbol{A}\,\boldsymbol{\Omega}))^2\;\tau^2\right]\,\boldsymbol{a}\right\} - \frac{\boldsymbol{a}\boldsymbol{\delta}'\,\boldsymbol{\Omega}\,\boldsymbol{\nu}}{\gamma}\,\tau\,, \tag{62}$$

so that

$$\pi^* = \tfrac{1}{\gamma}\left\{\boldsymbol{\Omega}^{-1}\,(\boldsymbol{\delta}\,\alpha + \boldsymbol{\nu}) + (\boldsymbol{\delta}'\boldsymbol{\Omega}\boldsymbol{\delta})\left[\tfrac{1}{2}\,\tau\,\left(\alpha - 2\frac{\boldsymbol{\delta}'\,\boldsymbol{\Omega}\,\boldsymbol{\nu}}{\boldsymbol{\delta}'\,\boldsymbol{\Omega}\,\boldsymbol{\delta}}\right) + \tfrac{1}{4}\operatorname{Tr}(\boldsymbol{A}\,\boldsymbol{\Omega})\,\tau^2\right]\,\boldsymbol{a}\right\}\,. \tag{63}$$

As a reminder, the first term is what you would expect from (a naïve modification of) the classical Merton problem since $(\boldsymbol{\delta}\,\alpha + \boldsymbol{\nu})$ are the drifts of the assets. The second term proportional to $\boldsymbol{a}$ represents the correction due to co-integration. As the above expression shows, the perturbation around the 'Merton' portfolio decays as the terminal date approaches. In all, the optimal strategy is affine in α even in this generalized problem of optimal investment in several co-integrated assets.

The explicit solution constructed here shows that it is indeed classical. Moreover, the optimal strategy is admissible. For a complete verification theorem, one needs to check the uniform integrability condition by checking that $\mathbb{E}[\exp\{-\gamma X_T^{\pi^*}\}]$ is indeed finite. Requiring this imposes constraints on the model parameters. The proof is essentially straight forward and follows along the lines of Tourin and Yan (2013) and Benth and Karlsen (2005).

4. Empirical performance of strategy

In this section we showcase how the strategy behaves in a three-asset case. The agent trades in a portfolio consisting of shares in GOOGL, FB and AMZN (Google, Facebook, and Amazon) and we use Nasdaq ITCH data to estimate the parameters of the model. Specifically, we employ midprices sampled every minute of the day on Nov 3, 2014 (the first trading day of November 2014) to estimate a VAR(1) model on the log-prices by performing the regression

$$\boldsymbol{Z}_n = \boldsymbol{b} + \boldsymbol{C}\,\boldsymbol{Z}_{n-1} + \boldsymbol{\varepsilon}_n\,, \tag{64}$$

where $\boldsymbol{Z}_n = \log(\boldsymbol{Y}_n)$, $\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \dots$ are assumed to be i.i.d. tri-variate standard normal random variables (with covariance matrix $\boldsymbol{\Omega}$), and the 3×1 vector $\boldsymbol{b}$ and 3×3 matrix $\boldsymbol{C}$ are estimated from a multivariate regression. The residuals then provide the variance-covariance matrix of $\boldsymbol{\varepsilon}_1$. Table 1 shows the results and the variance-covariance of these estimators are shown in Table 2.

Ticker	$\widehat{\boldsymbol{b}}$	$\widehat{\boldsymbol{C}}$			$\widehat{\boldsymbol{\Omega}}$ ($\times 10^{-7}$)		
GOOGL	0.45	0.092	-0.021	-0.008	1.97	0.78	0.67
FB	0.08	-0.011	0.014	0.016	0.78	3.57	0.95
AMZN	0.35	0.055	-0.011	0.009	0.67	0.95	2.79

Table 1: Calibrated VAR(1) model for GOOGL, FB and AMZN on Nov 3, 2014 using 1 minute data (Nasdaq ITCH).

In Figure 1 we show the in-sample path of the co-integrating factor α_t which, as expected, is a mean-reverting process. The dashed lines in the figure show the mean-reverting level (inner band), plus and minus two standard deviations of the co-integrating factor around the mean-reverting level (outer bands). Clearly, these short-lived deviations from the mean-reverting level are the main driver of the profits in the agent's strategy which we discuss below.

Next, we dimensionally reduce the problem to the case of a single co-integration factor by the following sequence of transformations:

	$\hat{\boldsymbol{b}}_1$	$\hat{\boldsymbol{b}}_2$	$\hat{\boldsymbol{b}}_3$	$\hat{\boldsymbol{C}}_{11}$	$\hat{\boldsymbol{C}}_{12}$	$\hat{\boldsymbol{C}}_{13}$	$\hat{\boldsymbol{C}}_{21}$	$\hat{\boldsymbol{C}}_{22}$	$\hat{\boldsymbol{C}}_{23}$	$\hat{\boldsymbol{C}}_{31}$	$\hat{\boldsymbol{C}}_{32}$	$\hat{\boldsymbol{C}}_{33}$
$\hat{\boldsymbol{b}}_1$	1206.26	480.57	410.17	-197.99	33.57	-16.89	-78.88	13.37	-6.73	-67.32	11.42	-5.74
$\hat{\boldsymbol{b}}_2$	480.57	2188.79	580.13	-78.88	13.37	-6.73	-359.26	60.91	-30.64	-95.22	16.15	-8.12
$\hat{\boldsymbol{b}}_3$	410.17	580.13	1709.29	-67.32	11.42	-5.74	-95.22	16.15	-8.12	-280.56	47.57	-23.93
$\hat{\boldsymbol{C}}_{11}$	-197.99	-78.88	-67.32	40.02	-8.62	-3.22	15.94	-3.43	-1.28	13.61	-2.93	-1.09
$\hat{\boldsymbol{C}}_{12}$	33.57	13.37	11.42	-8.62	3.15	1.30	-3.43	1.25	0.52	-2.93	1.07	0.44
$\hat{\boldsymbol{C}}_{13}$	-16.89	-6.73	-5.74	-3.22	1.30	5.53	-1.28	0.52	2.20	-1.09	0.44	1.88
$\hat{\boldsymbol{C}}_{21}$	-78.88	-359.26	-95.22	15.94	-3.43	-1.28	72.62	-15.63	-5.84	19.25	-4.14	-1.55
$\hat{\boldsymbol{C}}_{22}$	13.37	60.91	16.15	-3.43	1.25	0.52	-15.63	5.71	2.36	-4.14	1.51	0.63
$\hat{\boldsymbol{C}}_{23}$	-6.73	-30.64	-8.12	-1.28	0.52	2.20	-5.84	2.36	10.04	-1.55	0.63	2.66
$\hat{\boldsymbol{C}}_{31}$	-67.32	-95.22	-280.56	13.61	-2.93	-1.09	19.25	-4.14	-1.55	56.71	-12.21	-4.56
$\hat{\boldsymbol{C}}_{32}$	11.42	16.15	47.57	-2.93	1.07	0.44	-4.14	1.51	0.63	-12.21	4.46	1.84
$\hat{\boldsymbol{C}}_{33}$	-5.74	-8.12	-23.93	-1.09	0.44	1.88	-1.55	0.63	2.66	-4.56	1.84	7.84

Table 2: Variance-Covariance ($\times 10^{-5}$) of the estimates $\hat{\boldsymbol{b}}$ and $\boldsymbol{C}$ appearing in (64).

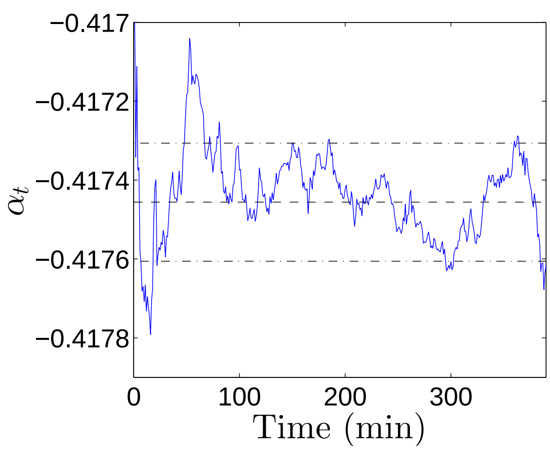

Figure 1: The in-sample in-sample path of the co-integrating factor α_t on Nov 3, 2014. The dashed lines show the mean reverting level (inner band) plus and minus two standard deviations of the co-integrating factor.

1. Define $\boldsymbol{\kappa} = \mathbb{I} - \boldsymbol{C}$ – the mean-reversion matrix.
2. Define $\boldsymbol{\Theta} = \boldsymbol{\kappa}^{-1}\,\boldsymbol{b}$ – the mean-reversion level.
3. Diagonalize $\boldsymbol{\kappa}$ by writing $\boldsymbol{\kappa} = \boldsymbol{U}\,\boldsymbol{D}\,\boldsymbol{U}^{-1}$ where $\boldsymbol{D}$ is the matrix of eigenvalues ordered largest to smallest, $\boldsymbol{U}$ is the corresponding matrix of eigenvectors stacked columnwise.
4. Define the dimensionally reduced matrix $\tilde{\boldsymbol{\kappa}} = \boldsymbol{U}\,\tilde{\boldsymbol{D}}\,\boldsymbol{U}^{-1}$, where $\tilde{\boldsymbol{D}} = \mathrm{diag}(\boldsymbol{D}_{11}, 0, 0, \dots, 0)$.
5. Then set $\boldsymbol{a}_k = -\tilde{\boldsymbol{\kappa}}_{1,k}$, $\boldsymbol{\delta}_k = \tilde{\boldsymbol{\kappa}}_{1,k}/\tilde{\boldsymbol{\kappa}}_{1,1}$ and $\boldsymbol{\nu}_k = (\tilde{\boldsymbol{\kappa}}\,\boldsymbol{\Theta})_k + \frac{1}{2}\boldsymbol{\Omega}_{kk}$.

The result of this is the set of parameters in the Table 3 below. Note that $\boldsymbol{\delta}'\,\boldsymbol{a} < 0$, so that the α process is indeed mean-reverting.

To show the performance of the strategy we run 10,000 simulations and record the profit & loss of each run. We use the parameters estimated above to simulate the short-term alpha component and co-integrated price paths (for every minute of the trading horizon) for Google, Facebook, and Amazon. For each run we assume that the agent starts with no asset holdings, all positions are liquidated at the end of the day, all transactions are at the midprice, and there are no transaction costs. In Figure 2 we show a single simulated sample path of α_t, the corresponding optimal position in the assets, the agent's wealth process for this path, and the

Ticker	$\boldsymbol{a}$	$\boldsymbol{\delta}$	$\boldsymbol{\nu}$
GOOGL	-0.0974	1.000	0.417
FB	0.0259	-0.004	-0.002
AMZN	0.0153	0.706	0.295

Table 3: The estimated model parameters as implied by the VAR(1) model estimated dimensionally reduced to one co-integration factor.

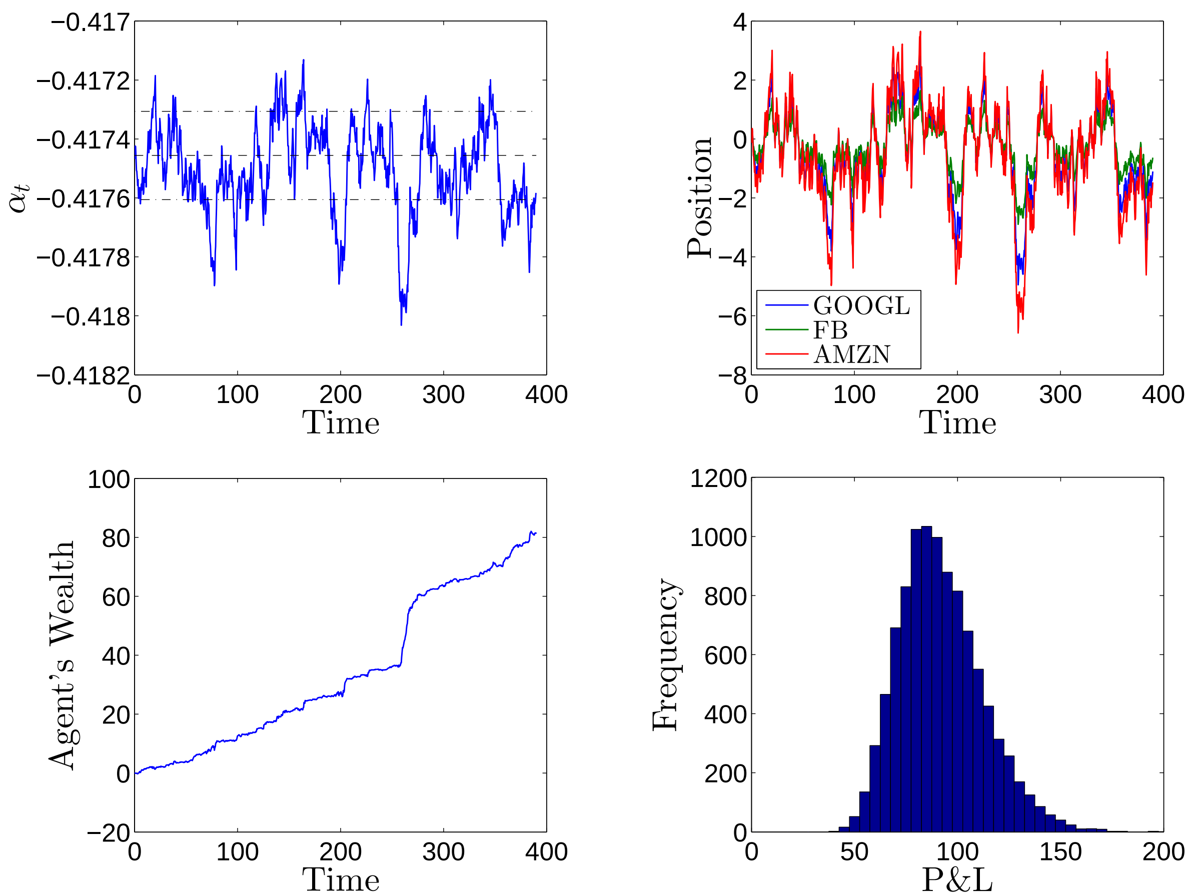

Figure 2: A simulated sample path of α_t (top left), the corresponding optimal position in the assets (top right), the agent's wealth process for this path (bottom left), and the histogram of the profit & loss over the 10,000 runs (bottom right).

histogram of the profit & loss over the 10,000 runs. The mean of the P&L is 91.99 and its standard deviation is 20.27.

For this particular run we observe that the agent's positions are always between 4 (long shares) and 6 (short shares) and the profit at the end of the trading horizon is \$80 approximately. The asset holdings exhibit a similar pattern to that of the short-term alpha, both depicted on the top panel in the figure. Recall that α_t enters (proportionally to δ_k) the drift of each asset, so times when α_t is above its mean-reverting level is when the midprices are expected to increase faster so the optimal strategy is to acquire the assets. Similarly, when α_t is below its mean-reverting level, is when growth in the midprices is slow (or could even be negative) so it is optimal to take short positions in the assets to benefit from the short-term drop in prices. Clearly, the effect of short-term alpha can be seen in the first term on the right-hand side of (63). In addition, recall that the structural dependence given by the co-integration factor, second term on the right-hand side of (63), is also playing a role in the optimal strategy.

5. Conclusions

We have shown how to optimally trade in and out of positions on a collection of co-integrated assets. Key to the strategy is how to model the structural dependence of asset prices through a common component in the drift of the asset returns. This common component is a co-integrating factor, which we label short-term alpha, and is the main source of profits in the optimal investment strategy.

In our model the investor takes positions in the assets to maximize expected utility of wealth. We have provided an explicit closed-form expression for the optimal (dynamic) investment strategy and have shown that it is affine in the co-integration factor. Moreover, we have shown that as the terminal date of the strategy approaches, the strategy behaves similar to a naïve modification of the Merton strategy where the constant drift (in the classical Merton model) is replaced by short-term alpha.

To illustrate the performance of the strategy we employ Nasdaq data to calibrate the model and use simulations to illustrate how the strategy performs when the investor takes positions in three assets: Google, Facebook, and Amazon.

Extensions to our work could look at i) how the strategy performs out-of-sample using realisations of the stock prices (instead of simulations), ii) how the strategy performs when the portfolio consists of liquid and illiquid assets, and iii) how the optimal investment strategy changes if the investor is ambiguity averse – i.e. recognises that the co-integration model might be misspecified and proceeds as in the work of Cartea et al. (2016).

References

Benth, F. E. and K. H. Karlsen (2005). A note on Merton's portfolio selection problem for the schwartz mean-reversion model. Stochastic analysis and applications 23(4), 687–704.

Cartea, Á., R. Donnelly, and S. Jaimungal (2016). Algorithmic trading with model uncertainty. SSRN: http://ssrn.com/abstract=2310645.

Cartea, Á., S. Jaimungal, and J. Penalva (2015). Algorithmic and High-Frequency Trading. Cambridge University Press.

Cartea, Á., S. Jaimungal, and J. Ricci (2014). Buy low, sell high: A high frequency trading perspective. SIAM Journal on Financial Mathematics 5(1), 415–444.

Duan, J.-C. and S. R. Pliska (2004). Option valuation with co-integrated asset prices. Journal of Economic Dynamics and Control 28(4), 727 – 754.

Elliott, R. J., J. Van Der Hoek, and W. P. Malcolm (2005). Pairs trading. Quantitative Finance 5(3), 271–276.

Lei, Y. and J. Xu (2015). Costly arbitrage through pairs trading. Journal of Economic Dynamics and Control 56, 1–19.

Leung, T. and X. Li (2015). Optimal mean reversion trading with transaction costs and stop-loss exit. International Journal of Theoretical and Applied Finance, 1550020.

Mudchanatongsuk, S., J. A. Primbs, and W. Wong (2008, June). Optimal pairs trading: A stochastic control approach. In 2008 American Control Conference, pp. 1035–1039. IEEE.

Ngo, M.-M. and H. Pham (2016). Optimal switching for pairs trading rule: a viscosity solutions approach. Journal of Mathematical Analysis and Applications 441(1), 403–425.

Tourin, A. and R. Yan (2013). Dynamic pairs trading using the stochastic control approach. Journal of Economic Dynamics and Control 37(10), 1972 – 1981.

Acknowledgments

We would like to thank two anonymous referees and the Editor. SJ would also like to thank NSERC and GRI for partially funding this work. The authors would like to thank Agnes Tourin for useful comments on an earlier version of this article.