Optimal Dynamic Pairs Trading of Futures Under a Two-Factor Mean-Reverting Model

Tim Leung* Raphael Yan†

August 8, 2018

Abstract

We study the problem of dynamically trading a pair of futures contracts. We consider a two-factor mean-reverting model, where the spot price tends to evolve around its stochastic equilibrium that is also mean-reverting. We derive the futures price dynamics and determine the optimal futures trading strategy by solving a utility maximization problem. By analyzing the associated Hamilton-Jacobi-Bellman equation, we solve the utility maximization explicitly and provide the optimal trading strategies in closed form. Our strategies are applied to volatility (VIX) futures trading, and illustrated in a series of numerical examples.

Keywords: dynamic trading, futures portfolio, mean-reverting model, utility maximization
JEL Classification: C61, D53, G11, G13
Mathematics Subject Classification (2010): 91G20, 91G80

*Department of Applied Mathematics, University of Washington, Seattle WA 98195. Email: timleung@uw.edu. Corresponding author.

†Department of Mathematics and Statistics, McMaster University, 1280 Main Street West, Hamilton, Ontario L8S 4K1, Canada. Email: raphaelyan1218@gmail.com

1 Introduction

A major class of quantitative trading strategies across many asset classes involves constructing portfolios of positions in two or more highly correlated, comoving, or cointegrated assets, such as stocks and exchange-traded funds, or derivatives, such as bonds and futures. In recent years, quantitative approaches have gained wide popularity among both retail and institutional investors thanks to the availability of vast market data, statistical models, and efficient computing algorithms.

One important example is pairs trading, where two related securities are traded, commonly in opposite positions. As a result, the trader hopes to capture profits from divergence and convergence while maintaining market-neutrality. A number of approaches apply machine learning and optimization algorithms to identify mean-reverting portfolios with a few assets from a larger collection of stocks (d'Aspremont, 2011; Zhang et al., 2018). Typically in pairs trading the portfolio is *static* during the trading horizon. There are numerical empirical studies on the empirical performance of pairs trading (Gatev et al., 2006) and timing of trades given mean-reverting prices (Elliott et al., 2005; Leung and Li, 2016, 2015; Kitapbayev and Leung, 2017).

In this paper, we combine the ideas of pairs trading and *dynamic* portfolio, and apply them to the trading of two futures with the same underlying. We consider a two-factor mean-reverting model, where the spot price tends to evolve around its stochastic equilibrium that is also mean-reverting. The model we work with is called the Central Tendency Ornstein Uhlenbeck (CTOU), as studied by Mencia and Sentana (2013) for pricing VIX futures. We determine the optimal futures trading strategy by solving a utility maximization problem. By analyzing the associated Hamilton-Jacobi-Bellman equation, we solve the utility maximization explicitly and provide the optimal trading strategies in closed form. Our strategies are applied to volatility (VIX) futures trading, and illustrated in a series of numerical examples.

Our work contributes to the literature on stochastic portfolio optimization. The seminal paper by Merton (1971) introduced stochastic control theory and Hamilton-Jacobi-Bellman (HJB) equation to portfolio optimization. Since then, there has been a wealth of research that incorporates more realistic or market-specific features into the optimization problem. For portfolio optimization when the assets are mean-reverting, Benth and Karlsen (2005), Boguslavsky and Boguslavskaya (2004), Simonsen (2003) or Wachter (2002) are early examples. Optimal stopping/switching approach, with features like transaction costs and stop-loss exits, are incorporated by Leung and Li (2015) and Pham and Ngo (2016). On the front of trading co-integrating assets, Tourin and Yan Tourin and Yan (2013) study a dynamic programming approach and provide analytical solutions and numerical results for the associated HJB equation. For more related studies, we refer to Leung and Li (2016) and references therein.

However, many assets individually are not mean-reverting, but a linear combination of them are. This gives rise to the statistical concept of cointegration initiated by Engle and Granger (1987), and the financial concept of pairs trading as described by Vidyamurthy (2004). From an economic perspective, Xiong (2001) solve for the equilibrium in a market populated with convergence traders, who are similar to traders practicing pairs-trading in our context, along with noise traders and long-term investors. A recent survey of the literature on statistical approaches to pairs trading is given in Krauss (2015).

On the other hand, futures has been an integral part of the global financial and continues to grow. There are also existing studies that investigate cointegration and trading strategies in the futures market. We refer to the representative paper by Brenner and Kroner (1995), as well as Dolatabadi et al. (2016) for a recent application in commodity futures. Leung et al. (2016) discuss the optimal timing to trade futures under three single-factor mean-reverting post models. Bichuch and Shreve (2013) consider trading a pair of futures but use the arithmetic Brownian motion and ignore the well-observed term-structure patterns in the futures market. Yamamoto and Hibiki (2017) examine large-scale multiple pairs-trading using a derivative-free optimization algorithm. Angoshtari (2016) provides theoretical conditions under which the pairs-trading optimization problem is market neutral. These related studies motivate us to consider a two-factor mean-reverting model to effectively capture the price dynamics of futures, and develop a stochastic control approach for pairs trading in the futures market.

The paper is structured as follows. We describe the two-factor mean-reverting spot model in Section 2. In Section 3, we discuss our portfolio optimization problem and examine some analytic properties. We provide illustrative numerical results from our model in Section 4, as well as concluding remarks in Section 5.

2 Futures Price Dynamics

The two-factor mean-reverting model we consider is called the Central Tendency Ornstein Uhlenbeck (CTOU). This model has been used for pricing volatility futures (see Mencia and Sentana (2013)). One major feature of this model is the mean-reverting dynamics of the spot price. Specifically, the spot price tends to evolve around its stochastic equilibrium, which is also mean-reverting. The CTOU is able to capture the stylistic features of empirically observed mean reversion in volatility indices and commodity prices. Empirically, the spot price mean-reverts relatively faster than the stochastic equilibrium to its long-run mean.

Moreover, our choice is also motivated by the model's tractability. As we will see, the structure of the associated stochastic differential equations (SDEs) is very amenable to analysis and allows us to obtain closed-form solutions for the optimization problem. As noted in Mencia and Sentana (2013), the simplicity of this model allows for easy estimation, and it is shown that the model fits well with historical data empirically. They further note that if a jump component is added, the resulting estimates become less stable, which suggests a jump component would unnecessarily complicate model estimation and application.

The spot price is denoted by V_t. The spot's log-price mean-reverts to a stochastic equilibrium process θ_t, which in turn mean-reverts to its own *constant* equilibrium level $\bar{\theta}$. Under the risk-neutral measure $\mathbb{Q}$, the log-price process and its stochastic equilibrium follow the SDEs

$$d\log V_t = \kappa(\theta_t - \log V_t)\,dt + \sigma\,d\tilde{Z}_t^v, \tag{1}$$

$$d\theta_t = \bar{\kappa}(\bar{\theta} - \theta_t)\,dt + \bar{\sigma}\,d\tilde{Z}_t^\theta. \tag{2}$$

Here, the constants κ and $\bar{\kappa}$ represent the speeds of mean reversion for $\log V_t$ and θ_t respectively, while σ and $\bar{\sigma}$ are the respective volatilities. The model has two independent standard Brownian motions, $\tilde{Z}_t^v$ and $\tilde{Z}_t^\theta$, defined on the probability space $(\Omega, \mathcal{F}, \mathbb{Q})$, with the filtration $\mathbb{F}$ generated by $(\tilde{Z}_t^v, \tilde{Z}_t^\theta)_{t\geq 0}$.

To related the dynamics of $\log V_t$ and θ_t to the physical measure $\mathbb{P}$, we specify the market prices of risk as the constants ζ and $\bar{\zeta}$. Under the physical measure $\mathbb{P}$, we have

$$dZ_t^v = d\tilde{Z}_t^v - \zeta\,dt, \tag{3}$$

$$dZ_t^\theta = d\tilde{Z}_t^\theta - \bar{\zeta}\,dt. \tag{4}$$

Note that the conditions under which the $\mathbb{P}$ measure is identical to $\mathbb{Q}$ is: $\zeta = \bar{\zeta} = 0$.

Thus we can recover the dynamics of V_t and θ_t under the physical measure $\mathbb{P}$ as

$$d\log V_t = \kappa\left(\theta_t + \frac{\sigma\zeta}{\kappa} - \log V_t\right)\,dt + \sigma\,dZ_t^v, \tag{5}$$

$$d\theta_t = \bar{\kappa}\left(\bar{\theta} + \frac{\bar{\sigma}\bar{\zeta}}{\bar{\kappa}} - \theta_t\right)\,dt + \bar{\sigma}\,dZ_t^\theta. \tag{6}$$

Remark 1 *The CTOU model is a variation of the concatenated SQR (CSQR)*

$$dV_t = \kappa(\theta_t - V_t)\,dt + \sigma\sqrt{V_t}\,d\tilde{Z}_t^v,$$
$$d\theta_t = \bar{\kappa}(\bar{\theta} - \theta_t)\,dt + \bar{\sigma}\sqrt{\theta_t}\,d\tilde{Z}_t^\theta,$$

which has been studied by Bates (2012), among others, and is presented in Mencia and Sentana (2013) as well.

We now consider futures contracts of different maturities written on the spot V. For the futures contract with maturity T_i, $i = 1, \dots, n$, we define the price at time $t \in [0, T_i]$ by

$$F^{(i)}(t, V_t, \theta_t) = \tilde{\mathbb{E}}[V_{T_i} \,|\, V_t, \theta_t].$$

We will continue to work with log prices for both the spot and its futures prices. Hence, we define

$$v_t \equiv \log V_t, \tag{7}$$

$$f^{(i)}(t, v_t, \theta_t) \equiv \log F^{(i)}(t, V_t, \theta_t). \tag{8}$$

From Appendix C of Mencia and Sentana (2013), we obtain the explicit log futures price:

$$\begin{aligned} f_t^{(i)} &\equiv f^{(i)}(t, v_t, \theta_t) \\ &= \bar{\theta} + D(T_i - t)(\theta_t - \bar{\theta}) + e^{-\kappa(T_i - t)}(\log V_t - \theta_t) + \frac{\sigma^2}{4\kappa}(1 - e^{-2\kappa(T_i - t)}) \\ &+ \frac{\bar{\sigma}^2}{2}\left(\frac{\kappa}{\kappa - \bar{\kappa}}\right)^2\left(\frac{1 - e^{-2\bar{\kappa}(T_i - t)}}{2\bar{\kappa}} + \frac{1 - e^{-2\kappa(T_i - t)}}{2\kappa} - 2\frac{1 - e^{-(\kappa + \bar{\kappa})(T_i - t)}}{\kappa + \bar{\kappa}}\right), \end{aligned}$$

where

$$D(\tau) = \frac{\kappa}{\kappa - \bar{\kappa}} e^{-\bar{\kappa}\tau} - \frac{\bar{\kappa}}{\kappa - \bar{\kappa}} e^{-\kappa\tau}.$$

Since our objective is to dynamically trade a futures portfolio under the physical measure $\mathbb{P}$, we derive the SDE for $f_t^{(i)}$ using Ito's Lemma (see Appendix A).

$$df_t^{(i)} = m_i(t)\,dt + \sigma e^{-\kappa(T_i - t)}\,dZ_t^v + \frac{\bar{\sigma}\kappa\big(e^{-\bar{\kappa}(T_i-t)} - e^{-\kappa(T_i-t)}\big)}{\kappa - \bar{\kappa}}\,dZ_t^\theta,$$

where the drift is a deterministic function time, given by

$$\begin{aligned} m_i(t) = {} & \frac{e^{-\bar{\kappa}(T_i-t)}}{\kappa - \bar{\kappa}}\kappa\bar{\sigma}\xi + \frac{e^{-\kappa(T_i-t)}}{(\kappa - \bar{\kappa})^2}(\kappa\bar{\sigma}\bar{\kappa}\xi - \kappa^2\bar{\sigma}\xi + \kappa^2\zeta\sigma - 2\kappa\zeta\bar{\kappa}\sigma + \zeta\bar{\kappa}^2\sigma) \\ & + \frac{\kappa^2\bar{\sigma}^2}{2(\kappa-\bar{\kappa})^2}\left(2e^{-(\kappa+\bar{\kappa})(T_i-t)} - e^{-2\bar{\kappa}(T_i-t)}\right) + \frac{e^{-2\kappa(T_i-t)}}{2(\kappa-\bar{\kappa})^2}(2\kappa\bar{\kappa}\sigma^2 - \kappa^2\bar{\sigma}^2 - \kappa^2\sigma^2 - \bar{\kappa}^2\sigma^2). \end{aligned} \tag{9}$$

In turn, we can write the futures price dynamics under $\mathbb{P}$ compactly as

$$\frac{dF_t^{(i)}}{F_t^{(i)}} = \mu_i(t)\,dt + \sigma_{vi}(t)\,dZ_t^v + \sigma_{\theta i}(t)\,dZ_t^\theta, \tag{10}$$

where all three time-deterministic coefficients are defined by

$$\sigma_{vi}(t) \equiv \sigma e^{-\kappa(T_i-t)}, \tag{11}$$

$$\sigma_{\theta i}(t) \equiv \frac{\bar{\sigma}\kappa\big(e^{-\bar{\kappa}(T_i-t)} - e^{-\kappa(T_i-t)}\big)}{\kappa - \bar{\kappa}}, \tag{12}$$

and

$$\mu_i(t) \equiv \frac{e^{-\bar{\kappa}(T_i-t)}\,\kappa\,\bar{\sigma}\,\bar{\zeta} - e^{-\kappa\,(T_i-t)}\,\left(\kappa\,\bar{\sigma}\,\bar{\zeta} - \kappa\,\zeta\,\sigma + \zeta\,\bar{\kappa}\,\sigma\right)}{\kappa - \bar{\kappa}} \tag{13}$$

$$= m_i(t) + \frac{\sigma_i(t)^2}{2}, \tag{14}$$

where

$$\sigma_i(t)^2 \equiv \sigma_{vi}(t)^2 + \sigma_{\theta i}(t)^2. \tag{15}$$

The instantaneous correlation between the two futures is defined by

$$\rho_{12}(t) \equiv \frac{\sigma_{v1}(t)\sigma_{v2}(t) + \sigma_{\theta 1}(t)\sigma_{\theta 2}(t)}{\sigma_1(t)\sigma_2(t)}, \tag{16}$$

which is a deterministic function of time only, independent of the state variables.

3 Utility Maximization Problem

Having established the dynamics of the futures prices in the previous section, we now consider the utility maximization problem involving a pair of futures. Let T_1 and T_2 be the maturities of the two futures in our portfolio. The optimization horizon will be denoted by T. Since the futures cannot be traded past expiry, we require $T_i \geq T$ for $i = 1, 2$. Using the futures price dynamics in (10), we write down the SDE for the portfolio wealth process as

$$dW_t = \pi_1(t, F_t^{(1)}, F_t^{(2)})\, dF_t^{(1)} + \pi_2(t, F_t^{(1)}, F_t^{(2)})\, dF_t^{(2)}, \tag{17}$$

where $\pi_i(t, F_t^{(1)}, F_t^{(2)})$, $i = 1, 2$, denote the number of contracts, and positive/negative values mean a long/short position, respectively. For brevity, we may write $\pi_i \equiv \pi_i(t, F_t^{(1)}, F_t^{(2)})$.

Re-writing in matrix form in terms of the two fundamental sources of randomness (Z_t^v, Z_t^θ), we get

$$\begin{bmatrix} dW_t \\ dF_t^{(1)} \\ dF_t^{(2)} \end{bmatrix} = \begin{bmatrix} \pi_1\mu_1(t)F_t^{(1)} + \pi_2\mu_2(t)F_t^{(2)} \\ \mu_1(t)F_t^{(1)} \\ \mu_2(t)F_t^{(2)} \end{bmatrix} dt$$

$$+ \begin{bmatrix} \pi_1\sigma_{v1}(t)F_t^{(1)} + \pi_2\sigma_{v2}(t)F_t^{(2)} & \pi_1\sigma_{\theta 1}(t)F_t^{(1)} + \pi_2\sigma_{\theta 2}(t)F_t^{(2)} \\ \sigma_{v1}(t)F_t^{(1)} & \sigma_{\theta 1}(t)F_t^{(1)} \\ \sigma_{v2}(t)F_t^{(2)} & \sigma_{\theta 2}(t)F_t^{(2)} \end{bmatrix} \begin{bmatrix} dZ_t^v \\ dZ_t^\theta \end{bmatrix}. \tag{18}$$

A pair of controls (π_1, π_2) is said to be admissible if (π_1, π_2) are real-valued, progressively measurable, and are such that the system of SDE (38) defines a unique solution $(W_t, F_t^{(1)}, F_t^{(2)})$ for every time $t \in [0, T]$ and $(\pi_1, \pi_2, F^{(1)}, F^{(2)})$ satisfy the integrability condition

$$\mathbb{E}\left(\int_t^T [\pi_1(s, F_s^{(1)}, F_s^{(2)})F_s^{(1)}]^2 + [\pi_2(s, F_s^{(1)}, F_s^{(2)})F_s^{(2)}]^2 ds\right) < \infty.$$

We denote by $\mathcal{A}_t$ the set of admissible controls with an initial time of investment t. Next, we define the value function $u(t, w, F_1, F_2)$ of the following optimization problem: the investor seeks an admissible strategy (π_1, π_2) that maximizes the utility from wealth at time T, that is,

$$u(t, w, F_1, F_2) = \sup_{(\pi_1, \pi_2) \in \mathcal{A}_t} \mathbb{E}\left(U(W_T) \,|\, W_t = w, F_t^{(1)} = F_1, F_t^{(2)} = F_2\right). \tag{19}$$

Here we only treat the case of the exponential utility function $U(w) = -e^{-\gamma w}$ where γ denotes the constant risk aversion coefficient.

3.1 HJB Equation and Closed-Form Solution

To facilitate presentation, we define the partial derivatives by

$$u_t = \frac{\partial u}{\partial t}, \quad u_w = \frac{\partial u}{\partial w}, \quad u_{ww} = \frac{\partial^2 u}{\partial w^2},$$

$$u_1 = \frac{\partial u}{\partial F_1}, \quad u_{11} = \frac{\partial^2 u}{\partial F_1^2}, \quad u_2 = \frac{\partial u}{\partial F_2}, \quad u_{22} = \frac{\partial^2 u}{\partial F_2^2},$$
$$u_{w1} = \frac{\partial^2 u}{\partial w \partial F_1}, \quad u_{w2} = \frac{\partial^2 u}{\partial w \partial F_2}, \quad u_{12} = \frac{\partial^2 u}{\partial F_1 \partial F_2}.$$

We determine the value function $u(t, w, F_1, F_2)$ by solving the HJB equation

$$\begin{aligned}
u_t + \sup_{\pi_1,\pi_2} \Bigg(& \pi_1 \mu_1(t) F_1 u_w + \pi_2 \mu_2(t) F_2 u_w \\
& + (\pi_1 \sigma_1(t)^2 F_1^2 + \pi_2(\sigma_{v1}(t)\sigma_{v2}(t) + \sigma_{\theta 1}(t)\sigma_{\theta 2}(t)) F_1 F_2) u_{w1} \\
& + (\pi_2 \sigma_2(t)^2 F_2^2 + \pi_1(\sigma_{v1}(t)\sigma_{v2}(t) + \sigma_{\theta 1}(t)\sigma_{\theta 2}(t)) F_1 F_2) u_{w2} \\
& + \frac{1}{2}(\pi_1^2 \sigma_1(t)^2 F_1^2 + \pi_2^2 \sigma_2(t)^2 F_2^2) u_{ww} \\
& + (\pi_1 \pi_2 (\sigma_{v1}(t)\sigma_{v2}(t) + \sigma_{\theta 1}(t)\sigma_{\theta 2}(t)) F_1 F_2) u_{ww} \Bigg) \\
& + \frac{\sigma_1(t)^2}{2} F_1^2 u_{11} + \frac{\sigma_2(t)^2}{2} F_2^2 u_{22} + \mu_1(t) F_1 u_1 + \mu_2(t) F_2 u_2 \\
& + (\sigma_{v1}(t)\sigma_{v2}(t) + \sigma_{\theta 1}(t)\sigma_{\theta 2}(t)) F_1 F_2 u_{12} = 0,
\end{aligned} \tag{20}$$

subject to the terminal condition

$$u(T, w, F_1, F_2) = -e^{-\gamma w}.$$

Next, we apply the transformation

$$u(t, w, F_1, F_2) = -e^{-\gamma w} G(t, f_1, f_2),$$

with $f_1 = \log F_1$ and $f_2 = \log F_2$. By direct substitution, we obtain the PDE for G:

$$\begin{aligned}
-e^{-\gamma w} G_t + \sup_{\pi_1,\pi_2} [\, & (\pi_1 \mu_1 F_1 + \pi_2 \mu_2 F_2) \gamma e^{-\gamma w} G \\
& + (\pi_1 \sigma_1^2 F_1^2 + \pi_2 \rho_{12} \sigma_1 \sigma_2 F_1 F_2) \gamma e^{-\gamma w} G_1 / F_1 + (\pi_2 \sigma_2^2 F_2^2 + \pi_1 \rho_{12} \sigma_1 \sigma_2 F_1 F_2) \gamma e^{-\gamma w} G_2 / F_2 \\
& + \frac{1}{2}(\pi_1^2 \sigma_1^2 F_1^2 + \pi_2^2 \sigma_2^2 F_2^2 + \rho_{12} \pi_1 \pi_2 \sigma_1 \sigma_2 F_1 F_2)(-\gamma^2 e^{-\gamma w} G) \,] \\
& + \mu_1 F_1 (-e^{-\gamma w} G_1 / F_1) + \mu_2 F_2 (-e^{-\gamma w} G_2 / F_2) \\
& + \frac{\sigma_1^2}{2} F_1^2 (e^{-\gamma w}(G_1 - G_{11})) / F_1^2 + \frac{\sigma_2^2}{2} F_2^2 (e^{-\gamma w}(G_2 - G_{22})) / F_2^2 \\
& + \rho_{12} \sigma_1 \sigma_2 F_1 F_2 (-e^{-\gamma w} G_{12}) / F_1 F_2 = 0,
\end{aligned} \tag{21}$$

where we have defined the partial derivatives

$$G_t = \frac{\partial G}{\partial t}, \quad G_1 = \frac{\partial G}{\partial f_1}, \quad G_2 = \frac{\partial G}{\partial f_2},$$

$$G_{11} = \frac{\partial^2 G}{\partial f_1^2}, \quad G_{22} = \frac{\partial^2 G}{\partial f_2^2}, \quad G_{12} = \frac{\partial^2 G}{\partial f_1 \partial f_2},$$

and suppressed the dependence on t, in μ_i, σ_i, σ_{vi}, $\sigma_{\theta i}$ and ρ_{12}, in order to simplify notation.

Cancelling $e^{-\gamma w}$ and rearranging, we get

$$\begin{aligned}
-G_t + \sup_{\pi_1,\pi_2}\Big[\quad & (\pi_1\mu_1 F_1 + \pi_2\mu_2 F_2)\gamma G \\
& + (\pi_1\sigma_1^2 F_1\gamma + \pi_2\gamma\rho_{12}\sigma_1\sigma_2 F_2)G_1 + (\pi_2\sigma_2^2 F_2\gamma + \pi_1\gamma\rho_{12}\sigma_1\sigma_2 F_1)G_2 \\
& - \frac{\sigma_1^2}{2}\pi_1^2 F_1^2\gamma^2 G - \frac{\sigma_2^2}{2}\pi_2^2 F_2^2\gamma^2 G - \gamma^2\pi_1\pi_2\rho_{12}\sigma_1\sigma_2 F_1 F_2 G \quad \Big] \\
& - \frac{\sigma_1^2}{2}(G_{11} - G_1) - \frac{\sigma_2^2}{2}(G_{22} - G_2) - \rho_{12}\sigma_1\sigma_2 G_{12} - \mu_1 G_1 - \mu_2 G_2 = 0,
\end{aligned} \tag{22}$$

with the terminal condition

$$G(T, f_1, f_2) = 1.$$

Performing the optimization in (22), we obtain the optimal controls

$$\pi_1^*(t, F_1, F_2) = \frac{\mu_1}{\gamma(1-\rho_{12}^2)\sigma_1^2 F_1} + \frac{G_1}{\gamma G F_1} - \rho_{12}\frac{\mu_2}{\gamma(1-\rho_{12}^2)\sigma_1\sigma_2 F_1}, \tag{23}$$

$$\pi_2^*(t, F_1, F_2) = \frac{\mu_2}{\gamma(1-\rho_{12}^2)\sigma_2^2 F_2} + \frac{G_2}{\gamma G F_2} - \rho_{12}\frac{\mu_1}{\gamma(1-\rho_{12}^2)\sigma_1\sigma_2 F_2}. \tag{24}$$

Then we substitute the optimal controls as in (23) and (24) to arrive at a *nonlinear* PDE for G:

$$\begin{aligned}
G_t = & \left(\frac{1}{2}\frac{\mu_1^2}{(1-\rho_{12}^2)\sigma_1^2} + \frac{1}{2}\frac{\mu_2^2}{(1-\rho_{12}^2)\sigma_2^2} - \frac{\rho_{12}\mu_1\mu_2}{(1-\rho_{12}^2)\sigma_1\sigma_2}\right) G \\
& + \frac{1}{2G}\left(G_1^2\sigma_1^2 + 2G_1 G_2\rho_{12}\sigma_1\sigma_2 + G_2^2\sigma_2^2\right) \\
& - \frac{1}{2}((G_{11} - G_1)\sigma_1^2 + (G_{22} - G_2)\sigma_2^2 + 2G_{12}\rho\sigma_1\sigma_2).
\end{aligned} \tag{25}$$

To solve (25), we apply another transformation

$$G(t, f_1, f_2) = e^{-\Phi(t, f_1, f_2)} \tag{26}$$

to (25) to obtain a *linear* PDE for Φ:

$$\begin{aligned}
0 = \Phi_t + & \left(\frac{1}{2}\frac{\mu_1^2}{(1-\rho_{12}^2)\sigma_1^2} + \frac{1}{2}\frac{\mu_2^2}{(1-\rho_{12}^2)\sigma_2^2} - \frac{\rho_{12}\mu_1\mu_2}{(1-\rho_{12}^2)\sigma_1\sigma_2}\right) \\
& + \frac{\sigma_1^2}{2}(\Phi_{11} - \Phi_1) + \frac{\sigma_2^2}{2}(\Phi_{22} - \Phi_2) + \rho_{12}\sigma_1\sigma_2\Phi_{12},
\end{aligned} \tag{27}$$

subject to $\Phi(T, f_1, f_2)$=0.

We can solve this linear PDE of Φ by using the ansatz

$$\Phi(t, f_1, f_2) = a_{11}(t)f_1^2 + a_1(t)f_1 + a_{22}(t)f_2^2 + a_2(t)f_2 + a_{12}(t)f_1 f_2 + a(t)$$

to deduce that

$$a'_{11}(t) = a'_{22}(t) = a'_{12}(t) = 0,\ a_{11}(t) = a_{22}(t) = a_{12}(t) = 0,$$
$$a'_1(t) = a'_2(t) = 0,\ a_1(t) = a_2(t) = 0.$$

From this, we deduce that Φ is a function of t only, independent of f_1 and f_2, and satisfies the first-order differential equation

$$\frac{d\Phi}{dt} = -\frac{\mu_1(t)^2\sigma_2(t)^2 + \mu_2(t)^2\sigma_1(t)^2 - 2\rho_{12}(t)\mu_1(t)\mu_2(t)\sigma_1(t)\sigma_2(t)}{2(1-\rho_{12}(t)^2)\sigma_1(t)^2\sigma_2(t)^2}.$$

Solving this and applying (13), (15), and (16), we obtain a closed-form expression for Φ. Precisely,

$$\Phi(t) = \frac{(T-t)\left(\zeta^2 + \bar{\zeta}^2\right)}{2}. \tag{28}$$

Unraveling the transformations, we write the value function as

$$u(t, w, F_1, F_2) = e^{-\gamma w + \Phi(t)} \tag{29}$$

Very interestingly, the value function depends on only two model parameters, namely, the market prices of risk ζ and $\bar{\zeta}$, along with the optimization horizon T. Moreover, the value function does not depend on the two futures current prices (F_1, F_2). The simplicity of the value function is unexpected, especially since there are two stochastic factors and two futures in the trading problem. This is a very useful result that shows clear dependence of the value function on the two model parameters (see Fig. 2 below). Nevertheless, it does not mean that the corresponding trading strategies are trivial. In fact, the strategies depend not only on other model parameters but also the futures prices, as we will discuss next.

3.2 Optimal Wealth Process

By applying (26) and (28) to (23) and (24), we obtain the optimal trading strategies (i.e. cash amounts invested in the futures)

$$\pi_1^*(t, F_1, F_2) = \frac{1}{\gamma(1-\rho_{12}(t)^2)\sigma_1(t)F_1}\left(\frac{\mu_1(t)}{\sigma_1(t)} - \rho_{12}(t)\frac{\mu_2(t)}{\sigma_2(t)}\right), \tag{30}$$
$$\pi_2^*(t, F_1, F_2) = \frac{1}{\gamma(1-\rho_{12}(t)^2)\sigma_2(t)F_2}\left(\frac{\mu_2(t)}{\sigma_2(t)} - \rho_{12}(t)\frac{\mu_1(t)}{\sigma_1(t)}\right). \tag{31}$$

We recall (13), (15), and (16), and express the optimal strategies explicitly in terms of model parameters. Precisely,

$$\pi_1^*(t, F_1, F_2) = \frac{-e^{-\bar{\kappa}(T_2-t)}\kappa\zeta\bar{\sigma} + e^{-\kappa(T_2-t)}\left(\kappa\zeta\bar{\sigma} + \kappa\bar{\zeta}\sigma - \bar{\kappa}\bar{\zeta}\sigma\right)}{e^{t(\kappa+\bar{\kappa})}\left(e^{-\bar{\kappa}T_1-\kappa T_2} - e^{-\kappa T_1-\bar{\kappa}T_2}\right)\kappa\gamma\bar{\sigma}\sigma F_1}, \tag{32}$$

$$\pi_2^*(t, F_1, F_2) = \frac{e^{-\bar{\kappa}(T_1-t)}\kappa\zeta\bar{\sigma} - e^{-\kappa(T_1-t)}\left(\kappa\zeta\bar{\sigma} + \kappa\bar{\zeta}\sigma - \bar{\kappa}\bar{\zeta}\sigma\right)}{e^{t(\kappa+\bar{\kappa})}\left(e^{-\bar{\kappa}T_1-\kappa T_2} - e^{-\kappa T_1-\bar{\kappa}T_2}\right)\kappa\gamma\bar{\sigma}\sigma F_2}. \tag{33}$$

Note that the optimal controls are functions of time and the futures prices, but not functions of the spot price V_t and its equilibrium level θ_t. In contrast to the well-known Merton portfolio under the exponential utility, the optimal strategies π_1^* and π_2^* here are independent of the trading horizon T or time t. The strategies are inversely proportional to γ, as is expected. For each $i \in \{1, 2\}$, the optimal strategy π_i^* depends only on the corresponding futures price F_i, but not the other futures price.

If we substitute the optimal controls π_1^* and π_2^* into the wealth process (17), we have

$$\begin{aligned} dW(t) &= \pi_1^*\, dF_t^{(1)} + \pi_2^*\, dF_t^{(2)} \\ &= \pi_1^* F_t^{(1)} \mu_1(t)\, dt + \pi_2^* F_t^{(2)} \mu_2(t)\, dt + \\ &+ \left(\pi_1^* \sigma_{v1}(t) F_t^{(1)} + \pi_2^* \sigma_{v2}(t) F_t^{(2)}\right) dZ_t^v + \left(\pi_1^* \sigma_{\theta 1}(t) F_t^{(1)} + \pi_2^* \sigma_{\theta 2}(t) F_t^{(2)}\right) dZ_t^\theta \\ &\equiv \mu_W\, dt + \sigma_W\, dZ_t^W, \end{aligned} \tag{34}$$

where we have defined

$$\begin{aligned} \mu_W &= \pi_1^* F_t^{(1)} \mu_1(t) + \pi_2^* F_t^{(2)} \mu_2(t), \\ \sigma_W^2 &= \left(\pi_1^* \sigma_{v1}(t) F_t^{(1)} + \pi_2^* \sigma_{v2}(t) F_t^{(2)}\right)^2 + \left(\pi_1^* \sigma_{\theta 1}(t) F_t^{(1)} + \pi_2^* \sigma_{\theta 2}(t) F_t^{(2)}\right)^2. \end{aligned}$$

Direct computation simplifies μ_W and σ_W to

$$\mu_W = \frac{\zeta^2 + \bar{\zeta}^2}{\gamma}, \qquad \text{and} \qquad \sigma_W^2 = \frac{\zeta^2 + \bar{\zeta}^2}{\gamma^2} = \frac{\mu_W}{\gamma}. \tag{35}$$

Substituting (35) into (34) implies that the optimal wealth process is in fact a Brownian motion with constant drift μ_W and volatility σ_W parameters. As a result, the optimal wealth depends on the market prices of risk, ζ and $\bar{\zeta}$ (see (3)-(4)), as well as the investor's risk aversion parameter γ.

4 Numerical Implementation

We will now further examine our results with numerical examples with simulated and empirical data. For our examples, we will set γ to be 1, and use the estimated parameters from the "full sample" in Table 4 of Mencia and Sentana (2013), which are displayed here in Table 1.

κ	$\bar{\kappa}$	$\bar{\theta}$	σ	$\bar{\sigma}$	ζ	$\bar{\zeta}$
5.827	0.300	3.019	1.037	0.446	−0.010	2.242

Table 1: CTOU model parameters

According to Section 3 in Mencia and Sentana (2013), the parameters are obtained from maximization of the so-called pseudo likelihood in state-space modeling, which is described in more details in Trolle and Schwartz (2009). The parameters so obtained were further tested by comparing to VIX option prices and compared on the basis of Root Mean Square Error (RMSE). As noted, one of the advantages of the CTOU model is its tractability, and in the context of estimation, the continuous time SDE for $\log V_t$ and θ_t can be easily written as a Gaussian VAR(1), for which the transition density are known in closed forms.

In Figure 1, we show the dependence of the optimal trading strategies, π_1^* and π_2^*, on the volatility $\bar{\sigma}$ of the stochastic equilibrium θ_t in the CTOU model. Observe that π_1^* is positive and increasing concave while π_1^* negative and decreasing convex. With the parameters given in Table 1, we are short the T_1-futures $F^{(1)}$ and long the T_2-futures $F^{(2)}$. When we rearrange the formulae (32) and (33) for π_1^* and π_2^*, and collect terms involving $\bar{\sigma}$, we see that for both $i = 1, 2$, the optimal strategies are of the form $A_i + B_i/\bar{\sigma}$, which means that the absolute value of the each strategy π_i^* decreases as $\bar{\sigma}$ gets large, with other variables held constant. The practical consequence is that the number of contracts held, on both the long and short sides, are decreasing as the volatility of the stochastic equilibrium increases. This is in line with a risk-averse trader's intuition, who would prefer less exposures on both legs of the paired-trade, if the volatility of the stochastic equilibrium is high.

Figure 2 illustrates how the optimal trading strategies, π_1^* and π_2^*, vary with respect to the time-to-maturity. We see the number of contracts to buy, or to sell, are both increasing as maturity increases, with π_1^* becoming more negative and π_2^* more positive. From the trader's perspective, this corresponds to taking bigger positions in the long end of the futures curve. As is well known (Alexander and Korovilas, 2013), the volatility of longer-term VIX contracts are in general lower than the short-term contracts. Therefore, under the CTOU model with parameters as calibrated to the VIX futures price history, it is optimal for the risk-averse trader to make larger bets on the pair trade when the contracts are far from maturity since volatility is lower in the long end as observed empirically.

In Figure 3 we compare the optimal trading strategies, π_1^* and π_2^* (for two futures) to the optimal strategy $\tilde{\pi}_i^*$ for trading a single futures. The case of dynamically trading a single futures is discussed in Appendix A.2 below. The optimal strategy is explicitly given in (37). As in Figure 2, we plot the strategies as functions of T_i, using same set of parameters. When trading a single contract, either with maturity T_1 or T_2, the corresponding optimal strategy, $\tilde{\pi}_1^*$ and $\tilde{\pi}_2^*$, is positive. This is in contrast to the two-contract case where the optimal strategies, π_1^* and π_2^*, are of opposite signs. This is intuitive because when two contracts are available, along with the fact that the two futures are based on the same sources of randomness, risk aversion drives the investor to reduce risks by taking long and short positions simultaneously.

A related question is: when only one contract is traded, does the investor favor the longer or shorter maturity? As we can see, $\tilde{\pi}_2^*$ is greater than $\tilde{\pi}_1^*$. This means that, given only one contract is available, the trader tends to take a larger position in the contract if it is further away from maturity. When we compare the long and short positions in the two-contract case to the single-contract case in terms of position size, we can see the optimal long-short strategy requires taking bigger positions in both contracts than either position in the single-contract case.

Using historical VIX futures data, we consider two contracts, one with maturity January 2011 and the other with maturity February 2011. We show the empirical optimal positions over the period October 2010 to December 2010. This period is chosen to correspond to the post-calibration period of the full sample in Table 4 of Mencia and Sentana (2013). Over this post-crisis period, the market was relatively calm compared to the market during the crisis, with the VIX index hovering around 20. Applying our the explicit formulae for the strategies, we compute π_1^*, π_2^*, and $\pi_1^* + \pi_2^*$ based on the daily settlement prices of these contracts as well as the parameters in Table 1. As shown in Figure 4, the optimal strategy π_1^* is negative throughout this period, corresponding to a short position in the front-month contract, and the opposite holds for π_2^*. However, the absolute value of position of π_2^* is larger, leading to a net positive position.

We now turn our attention to the value functions. To distinguish between the single-contract and two-contract cases, we let $\tilde{u}^{(i)}$ denote the value function in the single-contract case with the superscript (i) indicating the maturity T_i of that single contract in the portfolio. In Figure 5 we plot $\tilde{u}^{(1)}$, $\tilde{u}^{(2)}$ and u as functions of t, and set $w = 0$. We observe that the maximized expected utility from trading two contracts simultaneously is greater than the maximized expected utility derived from trading only a single contract regardless of the choice of maturity. In fact, the value function u is larger than the sum of the two value functions $\tilde{u}^{(1)}$ and $\tilde{u}^{(2)}$. This makes sense since the single-contract case can be viewed as two-contracts case but with one strategy constrained at zero. Effectively, the single-contract case is restricting the admissible set $\mathcal{A}_t$, thus reducing the maximum expected utility. Our result confirms the intuition that more choices of trading instruments are preferable to fewer.

Next, we consider the *certainty equivalent* for the trading opportunity in the two futures with wealth w at time t. Recall that the value function is in exponential form

$$u(t, w, F_1, F_2) = -e^{-\gamma w - \Phi(t)}.$$

We define the certainty equivalent by taking the inverse of the exponential utility function. Precisely, we have

$$C(t, w) = w + \frac{\Phi(t)}{\gamma}.$$

As we can see, the certainty equivalent is the sum of the investor's wealth w and the positive value $\frac{\Phi(t)}{\gamma}$. The latter is inversely proportional to the risk aversion parameter γ. Like the value function, the certainty equivalent does not depend on the current futures prices (F_1, F_2) but it does depend on the model parameters that drive the futures prices.

We now evaluate the behavior of C at time $t = 0$ and with zero initial wealth $W_0 = 0$. In other words, we will examine the following quantity:

$$C_0 = \frac{\Phi(0)}{\gamma},$$

and its sensitivity with respect to as we have plotted in Figure 6. In Figure 6 we plot the certainty equivalent against the price of risk. From (28) it is clear that C_0 is quadratic in ζ and $\bar{\zeta}$ under the CTOU model, and tends to infinity as the prices of risk increase.

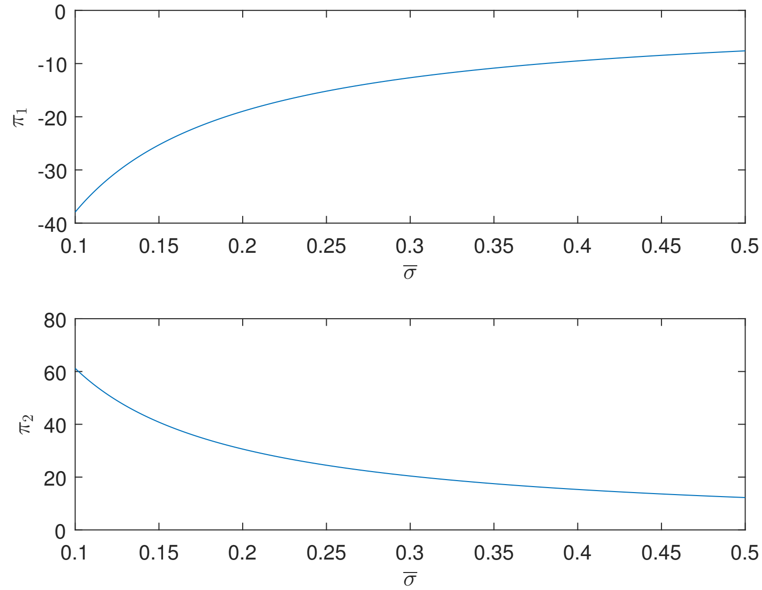

Figure 1: Optimal controls π_1^* and π_2^* as a function of $\bar{\sigma}$ under the CTOU model, with $\kappa = 5.827, \bar{\kappa} = 0.300, \bar{\theta} = 3.019, \sigma = 1.037, \zeta = -0.010$ and $\bar{\zeta} = 2.242$ as displayed in Table 1, at $T_1 = 30/365$ and $T_2 = 60/365$.

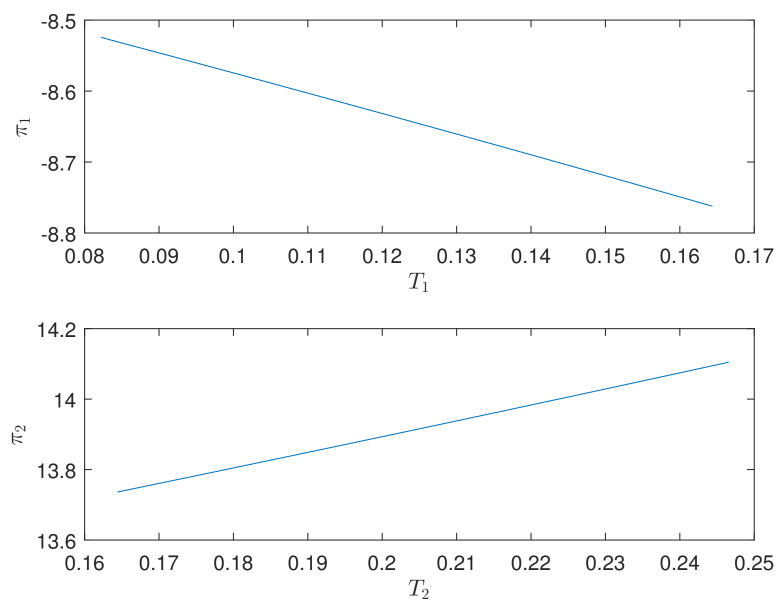

Figure 2: Optimal controls π_1^* and π_2^* as a function of T_1 and T_2 respectively, under the CTOU model, with parameters as displayed in Table 1, and T_1 ranges from [30/365,60/365], and T_2 ranges from [60/365,90/365].

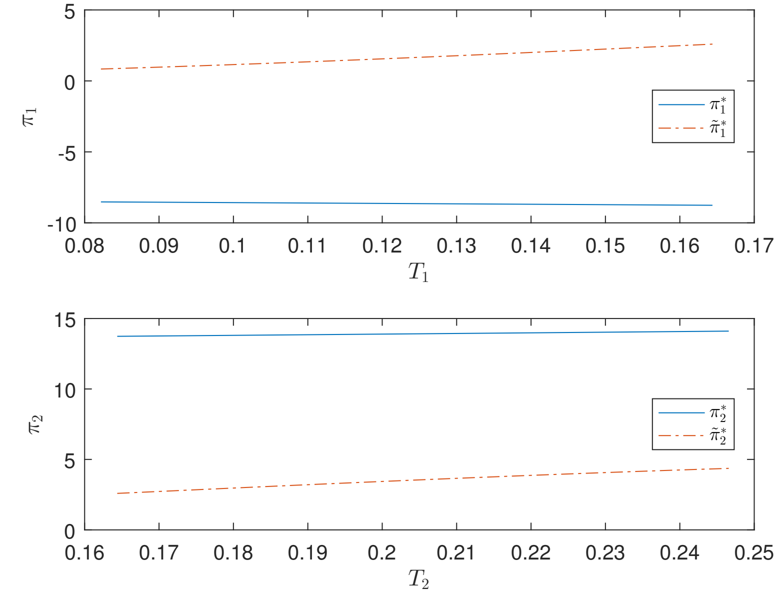

Figure 3: Optimal controls π_i^* with 2 contracts and $\tilde{\pi}_i^*$ with 1 contract, using parameters as displayed in Table 1, and T_1 ranges from [30/365,60/365], and T_2 ranges from [60/365,90/365].

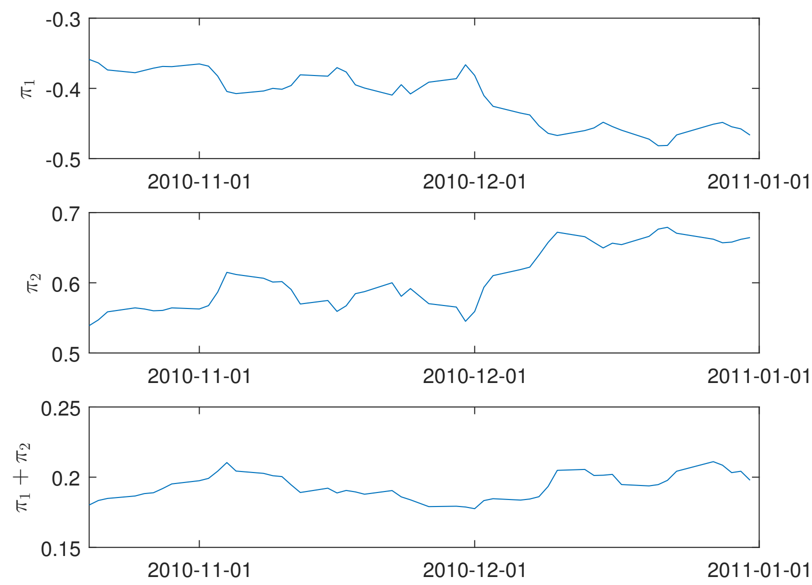

Figure 4: Optimal controls π_1^*, π_2^* and $\pi_1^* + \pi_2^*$ over the period Oct 2010-Dec 2010 using parameters as displayed in Table 1.

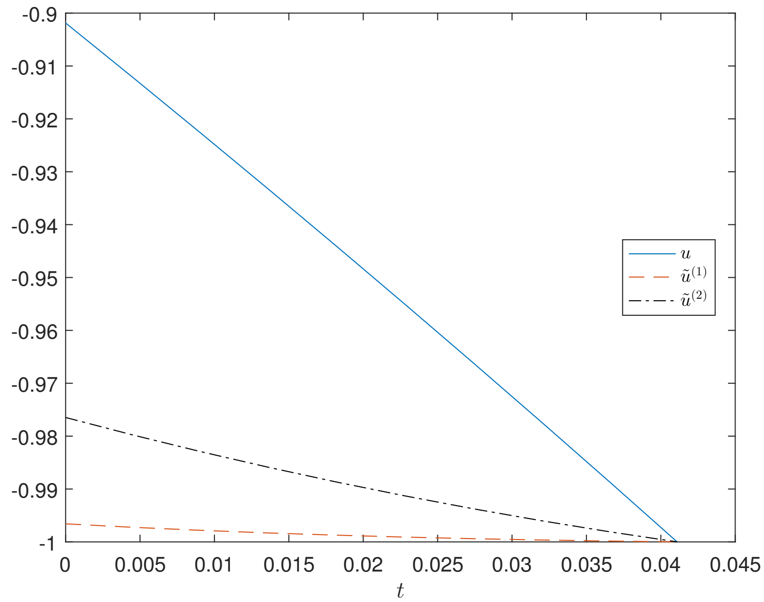

Figure 5: The value functions u, $\tilde{u}^{(1)}$ and $\tilde{u}^{(2)}$ at $w = 0$, with optimization horizon $T = 15/365$, maturity of F_1 is $T_1 = 30/365$, and maturity of $T_2 = 60/365$.

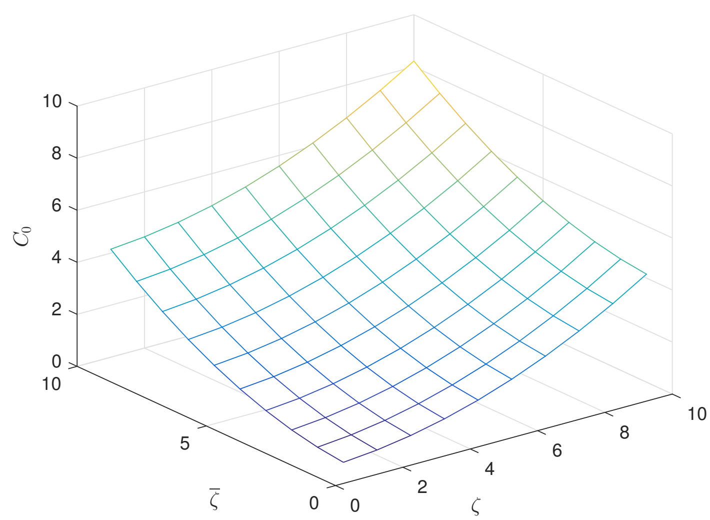

Figure 6: Certainty equivalent C_0 as a function of the market prices of risk ζ and $\bar{\zeta}$ under the CTOU model, with parameters as displayed in Table 1.

5 Conclusions

We have analyzed the problem of dynamically trading two futures contracts with the same underlying. Under a two-factor mean-reverting model for the spot price, we derive the futures price dynamics and solve the portfolio optimization problem in closed form and give explicit optimal trading strategies. By studying the associated Hamilton-Jacobi-Bellman equation, we solve the utility maximization explicitly and provide the optimal trading strategies in closed form. In addition to the analytic properties of our solutions, we also apply our results to VIX futures trading and present numerical examples to illustrate the optimal holdings.

There are a number of directions for future research. The method employed in this paper extends to a multidimensional setting higher than the 2 factor models considered here, such as a three-factor model with the addition of stochastic interest rate (see Schwartz (1997)). It will be also be interesting to incorporate other features such as stochastic volatility into the spot price dynamics, for example in the model described in Li and Tourin (2016), or jumps diffusion, as in other models described in Mencia and Sentana (2013). However, under more complicated models, the value functions and the optimal controls will likely require numerical approximations. Furthermore we will leave to future research to investigate the profitability of strategies based on our models using historical data.

A Appendix

A.1 Drift of $df_t^{(i)}$ under CTOU

By Ito's Lemma, the drift of $df_t^{(i)}$, denoted by $m_i(t)$, is given by

$$m_i(t) = \frac{df^{(i)}}{dt} + \kappa\left(\theta + \frac{\sigma\zeta}{\kappa} - v\right)\frac{df^{(i)}}{dv} + \bar{\kappa}\left(\bar{\theta} + \frac{\bar{\sigma}\bar{\zeta}}{\bar{\kappa}} - \theta\right)\frac{df^{(i)}}{d\theta}$$

We have the following derivatives

$$\begin{aligned}\frac{df^{(i)}(t,v,\theta)}{dt} &= e^{-\kappa(T_i-t)}\kappa\,(v-\theta) - \frac{\left(e^{\kappa T_i + t\,\bar{\kappa}} - e^{\kappa t + T_i\bar{\kappa}}\right)^2\kappa^2\bar{\sigma}^2}{2e^{2T_i(\kappa+\bar{\kappa})}(\kappa-\bar{\kappa})^2} \\ &\quad + \frac{\left(e^{-\kappa(T_i-t)} - e^{-\bar{\kappa}(T_i-t)}\right)\kappa\bar{\kappa}\left(\bar{\theta}-\theta\right)}{\kappa-\bar{\kappa}} - \frac{e^{-2\kappa(T_i-t)}\,\sigma^2}{2},\end{aligned}$$

$$\begin{aligned}&\kappa\left(\theta + \frac{\sigma\zeta}{\kappa} - v\right)\frac{df^{(i)}(t,v,\theta)}{dv} + \bar{\kappa}\left(\bar{\theta} + \frac{\bar{\sigma}\bar{\zeta}}{\bar{\kappa}} - \theta\right)\frac{df^{(i)}(t,v,\theta)}{d\theta} = \\ &\frac{e^{-(\kappa+\bar{\kappa})(T_i-t)}\left(e^{\kappa(T_i-t)}\kappa\left(\bar{\theta}\bar{\kappa} - \theta\bar{\kappa} + \bar{\sigma}\bar{\zeta}\right) + e^{\bar{\kappa}(T_i-t)}\left(\zeta\,(\kappa-\bar{\kappa})\,\sigma - \left(\kappa\left(\kappa\,(v-\theta) + (\bar{\theta}-v)\bar{\kappa} + \bar{\sigma}\bar{\zeta}\right)\right)\right)\right)}{\kappa-\bar{\kappa}}.\end{aligned}$$

In turn, we obtain

$$m_i(t) = \frac{e^{-\bar\kappa(T_i-t)}}{\kappa-\bar\kappa}\kappa\bar\sigma\xi + \frac{e^{-\kappa(T_i-t)}}{(\kappa-\bar\kappa)^2}(\kappa\bar\sigma\bar\kappa\xi - \kappa^2\bar\sigma\xi + \kappa^2\zeta\sigma - 2\kappa\zeta\bar\kappa\sigma + \zeta\bar\kappa^2\sigma)$$
$$+ \frac{\kappa^2\bar\sigma^2}{2(\kappa-\bar\kappa)^2}(2e^{-(\kappa+\bar\kappa)(T_i-t)} - e^{-2\bar\kappa(T_i-t)}) + \frac{e^{-2\kappa(T_i-t)}}{2(\kappa-\bar\kappa)^2}(2\kappa\bar\kappa\sigma^2 - \kappa^2\bar\sigma^2 - \kappa^2\sigma^2 - \bar\kappa^2\sigma^2).$$

Interestingly, the drift is a deterministic function of time, and does not depend on v_t, θ_t, and $\bar\theta$. To see this, we collect v, θ and $\bar\theta$ in $m_i(t)$, and get

$$\frac{df^{(i)}(t,v,\theta)}{dt} = e^{-\kappa\,(T_i-t)}\,\kappa\,v - \frac{\kappa^2\,\bar\sigma^2\left(e^{\kappa\,T_i+t\,\bar\kappa} - e^{\kappa\,t+T_i\,\bar\kappa}\right)^2}{2\,e^{2\,T_i\,(\kappa+\bar\kappa)}\left(\kappa-\bar\kappa\right)^2} + \frac{\bar\theta\,\kappa\,\bar\kappa\,\left(e^{-\kappa\,(T_i-t)} - e^{-\bar\kappa(T_i-t)}\right)}{\kappa-\bar\kappa}$$
$$+ \frac{\kappa\,\theta\,\left(e^{-\bar\kappa(T_i-t)}\,\bar\kappa - e^{-\kappa\,(T_i-t)}\,\kappa\right)}{\kappa-\bar\kappa} - \frac{e^{-2\,\kappa\,(T_i-t)}\,\sigma^2}{2},$$

$$\kappa\left(\theta + \frac{\sigma\zeta}{\kappa} - v\right)\frac{df^{(i)}(t,v,\theta)}{dv} = e^{-\kappa\,(T_i-t)}\,\kappa\,\theta + e^{-\kappa\,(T_i-t)}\,\left(\zeta\,\sigma - \kappa\,v\right),$$

$$\bar\kappa\left(\bar\theta + \frac{\bar\sigma\bar\zeta}{\bar\kappa} - \theta\right)\frac{df^{(i)}(t,v,\theta)}{d\theta} = \frac{\bar\theta\kappa\,\bar\kappa\left(e^{-\bar\kappa(T_i-t)} - e^{-\kappa\,(T_i-t)}\right)}{\kappa-\bar\kappa} + \frac{\theta\kappa\bar\kappa\,\left(e^{-\kappa\,(T_i-t)} - e^{-\bar\kappa(T_i-t)}\right)}{\kappa-\bar\kappa}$$
$$+ \frac{\kappa\,\bar\sigma\bar\zeta\left(e^{-\bar\kappa(T_i-t)} - e^{-\kappa\,(T_i-t)}\right)}{\kappa-\bar\kappa}.$$

When added together, the terms involving v, θ and $\bar\theta$ cancelled out, and we are left with (9).

A.2 Portfolio with a Single Futures Contract

We now discuss the case when the portfolio consists of only one futures contract (of maturity T_i). The system of SDEs for the wealth process and futures price is

$$\begin{bmatrix} dW_t \\ dF_t^{(1)} \end{bmatrix} = \begin{bmatrix} \tilde\pi_1\mu_1(t)F_t^{(1)} \\ \mu_1(t)F_t^{(1)} \end{bmatrix} dt + \begin{bmatrix} \tilde\pi_1\sigma_{v1}(t)F_t^{(1)} & \tilde\pi_1\sigma_{\theta1}(t)F_t^{(1)} \\ \sigma_{v1}(t)F_t^{(1)} & \sigma_{\theta1}(t)F_t^{(1)} \end{bmatrix} \begin{bmatrix} dZ_t^v \\ dZ_t^\theta \end{bmatrix}, \tag{36}$$

where we use the tilde notation to denote the single-contract case. In order to avoid confusion when we later compare the optimal controls and the value function to the two contracts case, we keep the subscript i in $\tilde\pi_i \equiv \tilde\pi_i(t, F_i)$ to denote the optimal control in the single contract case when the contract has a maturity of $T_i, i = 1, 2$.

We expect the value function $\tilde u(t, w, F_1)$ to solve the HJB equation

$$\tilde u_t + \sup_{\tilde\pi_1}[\,\tilde\pi_1\mu_1(t)F_1\tilde u_w + \tilde\pi_1\sigma_1(t)^2F_1^2\tilde u_{w1} + \frac{1}{2}\tilde\pi_1^2\sigma_1(t)^2F_1^2\tilde u_{ww}\,]$$
$$+ \frac{\sigma_1(t)^2}{2}F_1^2\tilde u_{11} + \mu_1(t)F_1\tilde u_1 = 0,$$

and the optimal control $\tilde{\pi}_1^*$ is given by

$$\tilde{\pi}_1^*(t, F_1) = \frac{\tilde{u}_w \mu_1(t) + F_1 \tilde{u}_{w1} \sigma_1(t)^2}{F_1 \tilde{u}_{ww} \sigma_1(t)^2}.$$

After substituting in $\tilde{\pi}_1^*$, we have the equation

$$\tilde{u}_t - \frac{\tilde{u}_w^2 \mu_1(t)^2}{2\tilde{u}_{ww}\sigma_1(t)^2} - \frac{F_1 \tilde{u}_w \tilde{u}_{w1} \mu_1(t)}{\tilde{u}_{ww}} + \frac{F_1(2\tilde{u}_1 \tilde{u}_{ww} \mu_1(t) - F_1(\tilde{u}_{w1}^2 - \tilde{u}_{11}\tilde{u}_{ww})\sigma_1(t)^2)}{2\tilde{u}_{ww}} = 0.$$

Next, we apply the transformation

$$\tilde{u}(t, w, F_1) = -e^{-\gamma w} e^{\tilde{\Phi}(t, f_1)},$$

where $f_1 = \log F_1$, and the tilde on Φ to again denote the single contract case, to get the PDE for $\tilde{\Phi}$ as

$$-\tilde{\Phi}_t = \frac{\mu_1(t)^2}{2\sigma_1(t)^2} + \frac{\sigma_1(t)^2}{2}(\tilde{\Phi}_{11} - \tilde{\Phi}_1),$$

subject to $\tilde{\Phi}(T, f_1)$=0. We see that $\tilde{\Phi}$ is a function of t only, and it satisfies

$$\frac{d\tilde{\Phi}}{dt} = -\frac{\mu_1(t)^2}{2\sigma_1(t)^2} = -\frac{\left(e^{-\bar{\kappa}\,(T_1-t)}\,\kappa\,\bar{\sigma}\,\xi - e^{-\kappa\,(T_1-t)}\,\left(\kappa\,\bar{\sigma}\,\xi - \kappa\,\zeta\,\sigma + \zeta\,\bar{\kappa}\,\sigma\right)\right)^2}{2\,(\kappa-\bar{\kappa})^2\,\left(\frac{\left(e^{-\kappa\,(T_1-t)} - e^{-\bar{\kappa}\,(T_1-t)}\right)^2 \kappa^2\,\bar{\sigma}^2}{(\kappa-\bar{\kappa})^2} + e^{-2\,\kappa\,(T_1-t)}\,\sigma^2\right)}.$$

In turn, we numerically evaluate the integral

$$\tilde{\Phi}(t) = \int_t^T \frac{\mu_1(t')^2}{2\sigma_1(t')^2} dt'.$$

Now if we express the optimal control in terms of $\tilde{\Phi}$, we will have

$$\tilde{\pi}_1^*(t, F_1) = \frac{\mu_1(t) - \sigma_1(t)^2 \tilde{\Phi}_1}{\gamma F_1 \sigma_1(t)^2} = \frac{\mu_1(t)}{\gamma F_1 \sigma_1(t)^2},$$

since $\tilde{\Phi}$ is a function of t only. We can see from (37) that $\tilde{\pi}_1^*$ is identical to that in the 2 contracts case, when $\rho_{12}(t)$ as defined in (16) equals zero. Explicitly, the optimal strategy $\tilde{\pi}_1^*$ in the single-contract case is given by

$$\tilde{\pi}_1^*(t, F_1) = \frac{-\left(e^{-\bar{\kappa}(T_1-t)}\,\bar{\sigma}\,\kappa\,\xi\right) + e^{-\kappa\,(T_1-t)}\,\left(\bar{\sigma}\,\kappa\,\xi - \zeta\,\kappa\,\sigma + \zeta\,\bar{\kappa}\,\sigma\right)}{\gamma F_1\,\left(\bar{\kappa}-\kappa\right)\,\left(\frac{\left(e^{-\kappa\,(T_1-t)} - e^{-\bar{\kappa}\,(T_1-t)}\right)^2 \bar{\sigma}^2\,\kappa^2}{(\kappa-\bar{\kappa})^2} + e^{-2\,\kappa\,(T_1-t)}\,\sigma^2\right)}. \tag{37}$$

A.3 Portfolio with Three Futures Contracts

Let us consider a portfolio of three futures contracts with different maturities T_1, T_2 and T_3. The wealth along with the futures prices follow the SDEs

$$\begin{bmatrix} dW_t \\ dF_t^{(1)} \\ dF_t^{(2)} \\ dF_t^{(3)} \end{bmatrix} = \begin{bmatrix} \pi_1\mu_1(t)F_t^{(1)} + \pi_2\mu_2(t)F_t^{(2)} + \pi_3\mu_3(t)F_t^{(3)} \\ \mu_1(t)F_t^{(1)} \\ \mu_2(t)F_t^{(2)} \\ \mu_3(t)F_t^{(3)} \end{bmatrix} dt$$

$$+ \begin{bmatrix} \pi_1\sigma_{v1}F_t^{(1)} + \pi_2\sigma_{v2}F_t^{(2)} + \pi_3\sigma_{v3}F_t^{(3)} & \pi_1\sigma_{\theta1}F_t^{(1)} + \pi_2\sigma_{\theta2}F_t^{(2)} + \pi_3\sigma_{\theta3}F_t^{(3)} \\ \sigma_{v1}F_t^{(1)} & \sigma_{\theta1}F_t^{(1)} \\ \sigma_{v2}F_t^{(2)} & \sigma_{\theta2}F_t^{(2)} \\ \sigma_{v3}F_t^{(3)} & \sigma_{\theta3}F_t^{(3)} \end{bmatrix} \begin{bmatrix} dZ_t^v \\ dZ_t^\theta \end{bmatrix}. \quad (38)$$

We expect the value function $u(t, w, F_1, F_2, F_3)$ to solve the HJB equation

$$\begin{aligned} u_t + \sup_{\pi_1,\pi_2,\pi_3} & [\, \pi_1\mu_1(t)F_1u_w + \pi_2\mu_2(t)F_2u_w + \pi_3\mu_2(t)F_3u_w \\ & + F_1(\pi_1\sigma_1(t)^2F_1 + \pi_2(\sigma_{v1}\sigma_{v2} + \sigma_{\theta1}\sigma_{\theta2})F_2 + \pi_3(\sigma_{v1}\sigma_{v3} + \sigma_{\theta1}\sigma_{\theta3})F_3)u_{w1} \\ & + F_2(\pi_1(\sigma_{v1}\sigma_{v2} + \sigma_{\theta1}\sigma_{\theta2})F_1 + \pi_2\sigma_2(t)^2F_2 + \pi_3(\sigma_{v2}\sigma_{v3} + \sigma_{\theta2}\sigma_{\theta3})F_3)u_{w2} \\ & + F_3(\pi_1(\sigma_{v1}\sigma_{v3} + \sigma_{\theta1}\sigma_{\theta3})F_1 + \pi_2(\sigma_{v2}\sigma_{v3} + \sigma_{\theta2}\sigma_{\theta3})F_2 + \pi_3\sigma_3(t)^2F_3)u_{w3} \\ & + ((F_1\pi_1\sigma_{v1} + F_2\pi_2\sigma_{v2} + F_3\pi_3\sigma_{v3})^2 + (F_1\pi_1\sigma_{\theta1} + F_2\pi_2\sigma_{\theta2} + F_3\pi_3\sigma_{\theta3})^2)u_{ww}\,] \\ & + \frac{\sigma_1(t)^2}{2}F_1^2u_{11} + \frac{\sigma_2(t)^2}{2}F_2^2u_{22} + \frac{\sigma_3(t)^2}{2}F_3^2u_{33} \\ & + \mu_1(t)F_1u_1 + \mu_2(t)F_2u_2 + \mu_3(t)F_3u_3 + (\sigma_{v1}\sigma_{v2} + \sigma_{\theta1}\sigma_{\theta2})F_1F_2u_{12} \\ & + (\sigma_{v1}\sigma_{v3} + \sigma_{\theta1}\sigma_{\theta3})F_1F_3u_{13} + (\sigma_{v2}\sigma_{v3} + \sigma_{\theta2}\sigma_{\theta3})F_2F_3u_{23} = 0. \end{aligned}$$

To solve for the optimal controls π_1, π_2, π_3, we impose the first-order conditions. This leads to the following system of equations:

$$u_{ww} \begin{bmatrix} F_1^2\sigma_1^2 & F_1F_2(\sigma_{v1}\sigma_{v2} + \sigma_{\theta1}\sigma_{\theta2}) & F_1F_3(\sigma_{v1}\sigma_{v3} + \sigma_{\theta1}\sigma_{\theta3}) \\ F_1F_2(\sigma_{v1}\sigma_{v2} + \sigma_{\theta1}\sigma_{\theta2}) & F_2^2\sigma_2^2 & F_2F_3(\sigma_{v2}\sigma_{v3} + \sigma_{\theta2}\sigma_{\theta3}) \\ F_1F_3(\sigma_{v1}\sigma_{v3} + \sigma_{\theta1}\sigma_{\theta3}) & F_2F_3(\sigma_{v2}\sigma_{v3} + \sigma_{\theta2}\sigma_{\theta3}) & F_3^2\sigma_3^2 \end{bmatrix} \begin{bmatrix} \pi_1 \\ \pi_2 \\ \pi_3 \end{bmatrix}$$

$$= - \begin{bmatrix} F_1u_w\mu_1 + F_1^2u_{w1}\sigma_1^2 + F_1F_2u_{w2}(\sigma_{v1}\sigma_{v2} + \sigma_{\theta1}\sigma_{\theta2}) + F_1F_3u_{w3}(\sigma_{v1}\sigma_{v3} + \sigma_{\theta1}\sigma_{\theta3}) \\ F_2u_w\mu_2 + F_1F_2u_{w1}(\sigma_{v1}\sigma_{v2} + \sigma_{\theta1}\sigma_{\theta2}) + F_2^2u_{w2}\sigma_2^2 + F_2F_3u_{w3}(\sigma_{v2}\sigma_{v3} + \sigma_{\theta2}\sigma_{\theta3}) \\ F_3u_w\mu_3 + F_1F_3u_{w1}(\sigma_{v1}\sigma_{v3} + \sigma_{\theta1}\sigma_{\theta3}) + F_2F_3u_{w2}(\sigma_{v2}\sigma_{v3} + \sigma_{\theta2}\sigma_{\theta3}) + F_3^2u_{w3}\sigma_3^2 \end{bmatrix}, \quad (39)$$

which is in fact singular.

References

Alexander, C. and Korovilas, D. (2013). Volatility Exchange-Traded Notes: Curse or Cure. *Journal of Alternative Investments*, 16(2):52–70.

Angoshtari, B. (2016). On the Market-Neutrality of Optimal Pairs-Trading Strategies. *ArXiv e-prints*.

Bates, D. S. (2012). U.S. stock market crash risk, 1926-2010. *Journal of Financial Economics*, 105:229–259.

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

Bichuch, M. and Shreve, S. (2013). Utility maximization trading two futures with transaction costs. *SIAM Journal of Financial Mathematics*, 4(1):26–85.

Boguslavsky, M. and Boguslavskaya, E. (2004). Arbitrage under power. *Risk*, pages 69–73.

Brenner, R. and Kroner, K. (1995). Arbitrage, cointegration, and testing the unbiasedness hypothesis in financial markets. *Journal of Financial and Quantitative Analysis*, 30(1):23–42.

d'Aspremont, A. (2011). Identifying small mean-reverting portfolios. *Quantitative Finance*, 11(3):351–364.

Dolatabadi, S., Nielsen, M. O., and Xu, K. (2016). A fractionally cointegrated VAR model with deterministic trends and application to commodity futures markets. *Journal of Empirical Finance*, 38(B):623–639.

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

Engle, R. and Granger, C. (1987). Co-integration and error correction: representation, estimation, and testing. *Econometrica*, 55(2):251–276.

Gatev, E., Goetzmann, W., and Rouwenhorst, K. (2006). Pairs trading: Performance of a relative-value arbitrage rule. *Review of Financial Studies*, 19(3):797–827.

Kitapbayev, Y. and Leung, T. (2017). Optimal mean-reverting spread trading: nonlinear integral equation approach. *Annals of Finance*, 13(2):181–203.

Krauss, C. (2015). Statistical arbitrage pairs trading strategies: Review and outlook. *IWQW Discussion Paper Series*, 9.

Leung, T., Li, J., Li, X., and Wang, Z. (2016). Speculative futures trading under mean reversion. *Asia-Pacific Financial Markets*, 23(4):281–304.

Leung, T. and Li, X. (2015). Optimal mean reversion trading with transaction costs and stop-loss exit. *International Journal of Theoretical & Applied Finance*, 18(3):15500.

Leung, T. and Li, X. (2016). *Optimal Mean Reversion Trading: Mathematical Analysis and Practical Applications*. Modern Trends in Financial Engineering. World Scientific, Singapore.

Li, T. N. and Tourin, A. (2016). Optimal pairs trading with time-varying volatility. *International Journal of Financial Engineering*, 3(3):1650023.

Mencia, J. and Sentana, E. (2013). Valuation of VIX derivatives. *Journal of Financial Economics*, 108(2):367–391.

Merton, R. (1971). Optimum consumption and portfolio rules in a continuous time model. *Journal of Economic Theory*, 3(4):373–413.

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

Schwartz, E. (1997). The stochastic behavior of commodity prices: Implications for valuation an hedging. *Journal of Finance*, 52(3):923–973.

Simonsen, K. (2003). Mean reversion and first passage times in relative value trading: the ornstein-uhlenbeck process. *ABN AMRO Fixed Income Relative Value Research*.

Tourin, A. and Yan, R. (2013). Dynamic pairs trading using the stochastic control approach. *Journal of Economic Dynamics and Control*, 37(10):1947–2156.

Trolle, A. and Schwartz, E. (2009). Unspanned stochastic volatility and the pricing of commodity derivatives. *The Review of Financial Studies*, 22(11):4423–4461.

Vidyamurthy, G. (2004). *Pairs Trading: Quantitative Methods and Analysis*. Wiley Finance.

Wachter, J. (2002). Portfolio and consumption decisions under mean-reverting returns: An exact solution for complete markets. *Journal of Financial and Quatitative Analysis*, 37(1):63–91.

Xiong, W. (2001). Convergence trading with wealth effects: an amplification mechanism in financial markets. *Journal of Financial Economics*, 62:247–292.

Yamamoto, R. and Hibiki, N. (2017). Optimal multiple pairs trading strategy using derivative free optimization under actual investment management conditions. *Journal of the Operations Research Society of Japan*, 60:244–261.

Zhang, J., Leung, T., and Aravkin, S. (2018). Sparse mean-reverting portfolios via penalized likelihood optimization. working paper.