

---

# Pairs Trading

## Under Geometric Brownian Motions

Q. Zhang with (R. Liu, Z. Wu, J. Tie and H. Zhang)

# What Is Pairs Trading?

---

# What Is Pairs Trading?

---

- Google Search: Pairs trading is a market-neutral trading strategy that matches a long position with a short position in a pair of highly correlated instruments such as two stocks, exchange-traded funds (ETFs), currencies, commodities or options.

# What Is Pairs Trading?

---

- Google Search: Pairs trading is a **market-neutral** trading strategy that matches a **long** position with a **short** position in a pair of **highly correlated** instruments such as two stocks, exchange-traded funds (ETFs), currencies, commodities or options.

# What Is Pairs Trading?

---

- Google Search: Pairs trading is a **market-neutral** trading strategy that matches a **long** position with a **short** position in a pair of **highly correlated** instruments such as two stocks, exchange-traded funds (ETFs), currencies, commodities or options.
  
- - Market-neutral:** Doesn't matter which direction the pair go
  - Long + short position:** Betting on their price difference
  - Highly correlated:** They need to be loosely 'tied' up

# What Is Pairs Trading?

- Google Search: Pairs trading is a **market-neutral** trading strategy that matches a **long** position with a **short** position in a pair of **highly correlated** instruments such as two stocks, exchange-traded funds (ETFs), currencies, commodities or options.

**Market-neutral:** Doesn't matter which direction the pair go

- **Long + short position:** Betting on their price difference

**Highly correlated:** They need to be loosely 'tied' up



# What Is Pairs Trading?

---

- Key: Simultaneously trade a pair of stocks with opposite directions.
- How: When their prices **diverge** (e.g., one stock moves up while the other moves down), the pairs trade would be triggered:  
Buy the weaker stock and short the stronger one and bet on the eventual price **convergence**.

# Why Pairs Trading?

---

- After all, we are not (neither our smart machines) that great forecasting market directions ...
- Investment strategies producing higher returns with smooth equity curve are highly desirable.
- Pairs trading is designed to address these issues and meet the needs.
- Major advantage: 'market neutral.'  
It can be profitable under any market (bull or bear) conditions.



# Implementation

---

It is important to determine when to **initiate** a pairs trade (i.e., how much divergence is sufficient to trigger a trade) and when to **close** the position (when to lock in profits).

# Brief Background and Literature Review

---

## Pairs Trading:

- Initially introduced by Bamberger and followed by Tartaglia's quantitative group at Morgan Stanley in the 1980s.
- Pairs trading (Gatev, Goetzmann, and Rouwenhorst, 2006)
- Book on pairs trading (Vidyamurthy, 2004)
- Pairs trading under a mean reversion model (Song and Zhang, 2013)
- Mean reversion trading (Zhang and Zhang, 2009 and Tie and Zhang 2016)

# What's New?

---

- One key element in traditional pairs trading studies is the **mean reversion** assumption (a difference of the prices is mean reversion).
- In order to meet the mean-reversion requirement, tradable pairs are often limited to the same industrial sector.
- From a practical viewpoint, it is highly desirable to have a broader range of stock selections for pairs trading.
- It would be interesting to study pairs trading under various other models to include, for example, **geometric Brownian motions**. Successful attempts would also put pairs trading practice on a firmer theoretical ground.
- Question: Would pairs trading work with GBM stocks?

# What's New?

---

- One key element in traditional pairs trading studies is the **mean reversion** assumption (a difference of the prices is mean reversion).
- In order to meet the mean-reversion requirement, tradable pairs are often limited to the same industrial sector.
- From a practical viewpoint, it is highly desirable to have a broader range of stock selections for pairs trading.
- It would be interesting to study pairs trading under various other models to include, for example, **geometric Brownian motions**. Successful attempts would also put pairs trading practice on a firmer theoretical ground.
- Question: Would pairs trading work with GBM stocks?
- Answer: Yes

# McDonald and Siegel's Optimal Stopping

---

- Back in 1986, McDonald and Siegel studied an optimal timing of investment in an irreversible project.
- Then in 1998, this problem was studied under precise optimality conditions by Hu and Øksendal. They also provided rigorous mathematical proofs.

## The Connection ...

---

- McDonald and Siegel's problem can be easily interpreted in terms of pairs trading.

## The Connection ...

---

- McDonald and Siegel's problem can be easily interpreted in terms of pairs trading.
- It is a simple pairs trading **selling** rule!

## How Different ...

---

- We extend their results by allowing **sequential** and simultaneous trading of these pairs.
- Focus:
  - (a) simple and easily implementable pairs trading strategy,
  - (b) closed-form solution, and
  - (c) its optimality.



# Our Focus

---

- A pair position: a **long** position in stock 1 and a **short** position in stock 2.
- Objective: To open (buy) and close (sell) these pairs positions sequentially to maximize a discounted payoff.
- A fixed percentage transaction cost will be imposed to each trade.
- Simple strategy: Determined by two threshold lines.
- (Dependence of these threshold levels on various parameters in numerical examples will be considered.)
- Implementation of the results with a pair of stocks based on their historical prices will be presented.

## By-product

---

In control applications, closed-form solutions for an optimal control problem are rare and very difficult to obtain.

We obtain a closed-form solution for an optimal decision making problem (nonlinear, two-dimensional, second order HJB equations with variational inequalities).

# The Model

---

We consider two stocks  $\mathbf{S}^1$  and  $\mathbf{S}^2$ . Let  $\{X_t^1, t \geq 0\}$  denote the prices of stock  $\mathbf{S}^1$  and  $\{X_t^2, t \geq 0\}$  that of stock  $\mathbf{S}^2$ . They satisfy the following stochastic differential equation:

$$d \begin{pmatrix} X_t^1 \\ X_t^2 \end{pmatrix} = \begin{pmatrix} X_t^1 & \\ & X_t^2 \end{pmatrix} \left[ \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} dt + \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix} d \begin{pmatrix} W_t^1 \\ W_t^2 \end{pmatrix} \right],$$

where  $\mu_i$ ,  $i = 1, 2$ , are the return rates,  $\sigma_{ij}$ ,  $i, j = 1, 2$ , the volatility constants, and  $(W_t^1, W_t^2)$  a 2-dimensional standard Brownian motion.

# The Model

---

- Assume the corresponding pairs position consists of one-share long position in stock  $S^1$  and one-share short position in stock  $S^2$ .
- Let  $Z$  denote the corresponding pairs position.
- One share in  $Z$  means the combination of one share long position in  $S^1$  and one share short position in  $S^2$ .
- The net position at any time can be either long (with one share of  $Z$ ) or flat (no stock position of either  $S^1$  or  $S^2$ ).

# The Model

---

- Let  $i = 0, 1$  denote the initial net position and let  $\tau_0 \leq \tau_1 \leq \tau_2 \leq \dots$  denote a sequence of stopping times.
- If initially the net position is long ( $i = 1$ ), then one should sell **Z** before acquiring any future shares.  
The corresponding trading sequence is denoted by  $\Lambda_1 = (\tau_0, \tau_1, \tau_2, \dots)$ .
- Likewise, if initially the net position is flat ( $i = 0$ ), then one should start to buy a share of **Z**.  
The corresponding sequence of stopping times is denoted by  $\Lambda_0 = (\tau_1, \tau_2, \dots)$ .

# The Model

---

Let  $K$  denote the transaction cost percentage (e.g., slippage and/or commission) associated with buying or selling of stocks  $\mathbf{S}^i$ ,  $i = 1, 2$ . For example, the cost to establish the pairs position  $\mathbf{Z}$  at  $t = t_1$  is

$$(1 + K)X_{t_1}^1 - (1 - K)X_{t_1}^2$$

and the proceeds to close it at a later time  $t = t_2$  is

$$(1 - K)X_{t_2}^1 - (1 + K)X_{t_2}^2.$$

For ease of notation, let  $\beta_b = 1 + K$  and  $\beta_s = 1 - K$ .

# The Model

Reward functions:  $J_0 = J_0(x_1, x_2, \Lambda_0)$  and  $J_1 = J_1(x_1, x_2, \Lambda_1)$

$$J_0 = E \left\{ \begin{aligned} &[e^{-\rho\tau_2}(\beta_s X_{\tau_2}^1 - \beta_b X_{\tau_2}^2)I_{\{\tau_2 < \infty\}} - e^{-\rho\tau_1}(\beta_b X_{\tau_1}^1 - \beta_s X_{\tau_1}^2)I_{\{\tau_1 < \infty\}}] \\ &+ [e^{-\rho\tau_4}(\beta_s X_{\tau_4}^1 - \beta_b X_{\tau_4}^2)I_{\{\tau_4 < \infty\}} - e^{-\rho\tau_3}(\beta_b X_{\tau_3}^1 - \beta_s X_{\tau_3}^2)I_{\{\tau_3 < \infty\}}] + \cdots \end{aligned} \right\},$$
$$J_1 = E \left\{ \begin{aligned} &e^{-\rho\tau_0}(\beta_s X_{\tau_0}^1 - \beta_b X_{\tau_0}^2)I_{\{\tau_0 < \infty\}} \\ &+ [e^{-\rho\tau_2}(\beta_s X_{\tau_2}^1 - \beta_b X_{\tau_2}^2)I_{\{\tau_2 < \infty\}} - e^{-\rho\tau_1}(\beta_b X_{\tau_1}^1 - \beta_s X_{\tau_1}^2)I_{\{\tau_1 < \infty\}}] \\ &+ [e^{-\rho\tau_4}(\beta_s X_{\tau_4}^1 - \beta_b X_{\tau_4}^2)I_{\{\tau_4 < \infty\}} - e^{-\rho\tau_3}(\beta_b X_{\tau_3}^1 - \beta_s X_{\tau_3}^2)I_{\{\tau_3 < \infty\}}] + \cdots \end{aligned} \right\}.$$

**Assumptions.**  $\rho > \mu_1$  and  $\rho > \mu_2$ .

Value functions:  $V_i(x_1, x_2) = \sup_{\Lambda_i} J_i(x_1, x_2, \Lambda_i)$ ,  $i = 0, 1$ .

## Example.

---

We consider stock prices of Target Corp. (TGT) and Wal-Mart Stores Inc. (WMT).

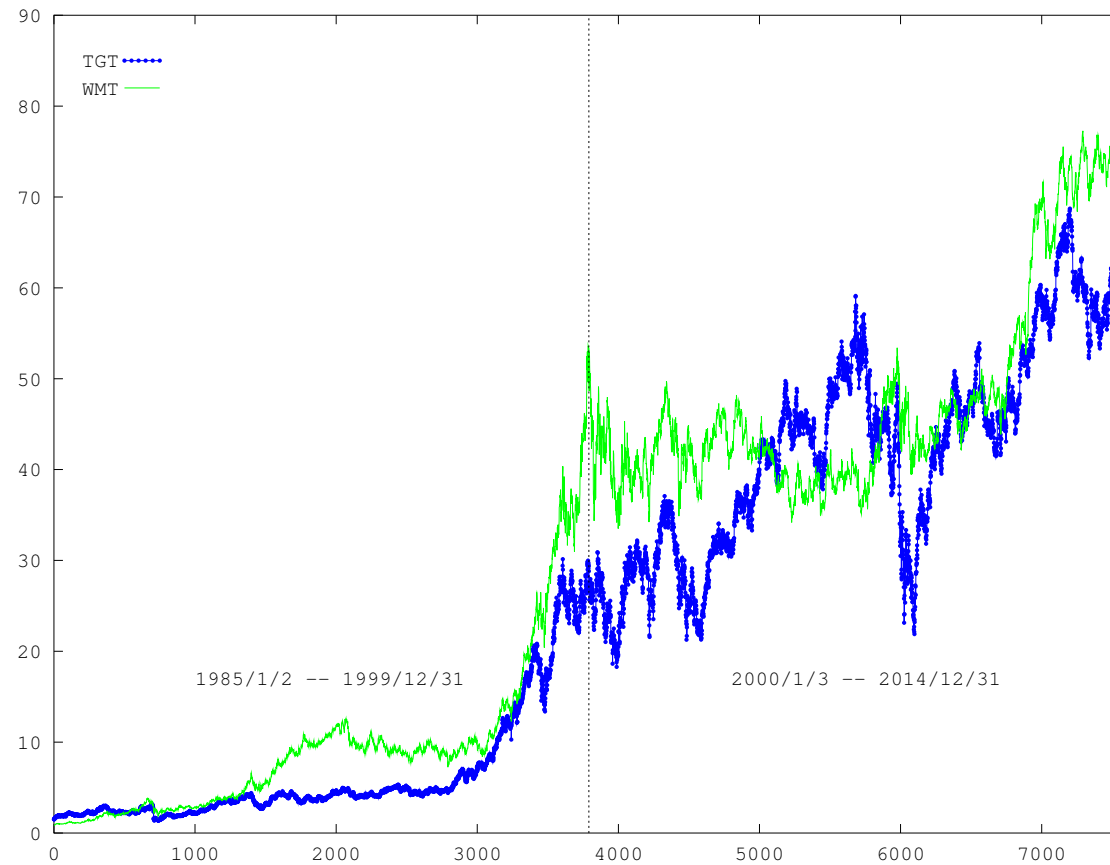
Daily closing prices of both stocks from 1985 to 2014 are divided into two sections:

Part 1 (1985-1999) is used to calibrate the model;

Part 2 (2000-2014) to backtest the performance of our results.



## Example: TGT and WMT Daily Closing Prices (1985 – 2014)



$$\mu_1 = 0.2059, \mu_2 = 0.2459, \sigma_{11} = 0.3112, \sigma_{12} = 0.0729, \sigma_{21} = 0.0729, \sigma_{22} = 0.2943.$$

# Value Function Properties

Given stopping times  $\tau_0 \leq \tau_1 \leq \tau_2, \dots$ , let  $\Lambda_1 = (\tau_0, \tau_1, \tau_2, \dots)$  and  $\Lambda_0 = (\tau_1, \tau_2, \dots)$ . Note that

$$J_1(x_1, x_2, \Lambda_1) = E[e^{-\rho\tau_0}(\beta_s X_{\tau_0}^1 - \beta_b X_{\tau_0}^2)I_{\{\tau_0 < \infty\}}] + J_0(x_1, x_2, \Lambda_0).$$

In particular, if  $\tau_0 = 0$ , a.s., then

$J_1(x_1, x_2, \Lambda_1) = \beta_s x_1 - \beta_b x_2 + J_0(x_1, x_2, \Lambda_0)$ . It follows that

$$\boxed{V_1(x_1, x_2) \geq \beta_s x_1 - \beta_b x_2 + V_0(x_1, x_2).}$$

Similarly, let  $\Lambda_0 = (\tau_1, \tau_2, \dots)$  and the subsequent  $\Lambda_1 = (\tau_2, \dots)$ . Then,

$$J_0(x_1, x_2, \Lambda_0) = -E[e^{-\rho\tau_1}(\beta_b X_{\tau_1}^1 - \beta_s X_{\tau_1}^2)I_{\{\tau_1 < \infty\}}] + J_1(x_1, x_2, \Lambda_1).$$

Setting  $\tau_1 = 0$ , a.s., leads to

$$\boxed{V_0(x_1, x_2) \geq -\beta_b x_1 + \beta_s x_2 + V_1(x_1, x_2).}$$

## Lower and Upper Bounds for $V_i(x_1, x_2)$

---

**Lemma.** For all  $x_1, x_2 > 0$ , we have

$$\begin{aligned} 0 &\leq V_0(x_1, x_2) \leq x_2, \\ \beta_s x_1 - \beta_b x_2 &\leq V_1(x_1, x_2) \leq \beta_b x_1 + K x_2. \end{aligned}$$

# HJB Equations

Let

$$\mathcal{A} = \frac{1}{2} \left\{ a_{11}x_1^2 \frac{\partial^2}{\partial x_1^2} + 2a_{12}x_1x_2 \frac{\partial^2}{\partial x_1 \partial x_2} + a_{22}x_2^2 \frac{\partial^2}{\partial x_2^2} \right\} + \mu_1x_1 \frac{\partial}{\partial x_1} + \mu_2x_2 \frac{\partial}{\partial x_2}$$

where

$$\begin{cases} a_{11} = \sigma_{11}^2 + \sigma_{12}^2, \\ a_{12} = \sigma_{11}\sigma_{21} + \sigma_{12}\sigma_{22}, \\ a_{22} = \sigma_{21}^2 + \sigma_{22}^2. \end{cases}$$

HJB equations: For  $(x_1, x_2) > 0$ ,

$$\begin{cases} \min \left\{ \rho v_0(x_1, x_2) - \mathcal{A}v_0(x_1, x_2), v_0(x_1, x_2) - v_1(x_1, x_2) + \beta_b x_1 - \beta_s x_2 \right\} = 0, \\ \min \left\{ \rho v_1(x_1, x_2) - \mathcal{A}v_1(x_1, x_2), v_1(x_1, x_2) - v_0(x_1, x_2) - \beta_s x_1 + \beta_b x_2 \right\} = 0. \end{cases}$$

## Solving $(\rho - \mathcal{A})v_i(x_1, x_2) = 0$

To solve  $(\rho - \mathcal{A})v_i(x_1, x_2) = 0$ ,  $i = 0, 1$ , consider

$$v_i(x_1, x_2) = c_{i1}x_1^{\delta_1}x_2^{1-\delta_1} + c_{i2}x_1^{-\delta_2}x_2^{1+\delta_2},$$

where  $\delta_1$  and  $(-\delta_2)$  are roots of the following equation:

$$\frac{1}{2}[a_{11}\delta(\delta - 1) + 2a_{12}\delta(1 - \delta) + a_{22}(1 - \delta)(-\delta)] + \mu_1\delta + \mu_2(1 - \delta) - \rho = 0.$$

Let  $\lambda = (a_{11} - 2a_{12} + a_{22})/2$ . We can write

$$\begin{aligned}\delta_1 &= \frac{1}{2} \left( 1 + \frac{\mu_2 - \mu_1}{\lambda} + \sqrt{\left( 1 + \frac{\mu_2 - \mu_1}{\lambda} \right)^2 + \frac{4\rho - 4\mu_2}{\lambda}} \right) > 1, \\ \delta_2 &= \frac{1}{2} \left( -1 - \frac{\mu_2 - \mu_1}{\lambda} + \sqrt{\left( 1 + \frac{\mu_2 - \mu_1}{\lambda} \right)^2 + \frac{4\rho - 4\mu_2}{\lambda}} \right) > 0.\end{aligned}$$

# Regions for the Variational Inequalities

---

Intuitively, if  $X_t^1$  is small and  $X_t^2$  is large, then one should buy  $S^1$  and sell (short)  $S^2$ . That is to open a pairs position  $\mathbf{Z}$ . If, on the other hand,  $X_t^1$  is large and  $X_t^2$  is small, then one should close the position  $\mathbf{Z}$  by selling  $S^1$  and buying back  $S^2$ .

In view of this, we divide the first quadrant

$P = \{(x_1, x_2) : x_1 > 0 \text{ and } x_2 > 0\}$  into three regions:

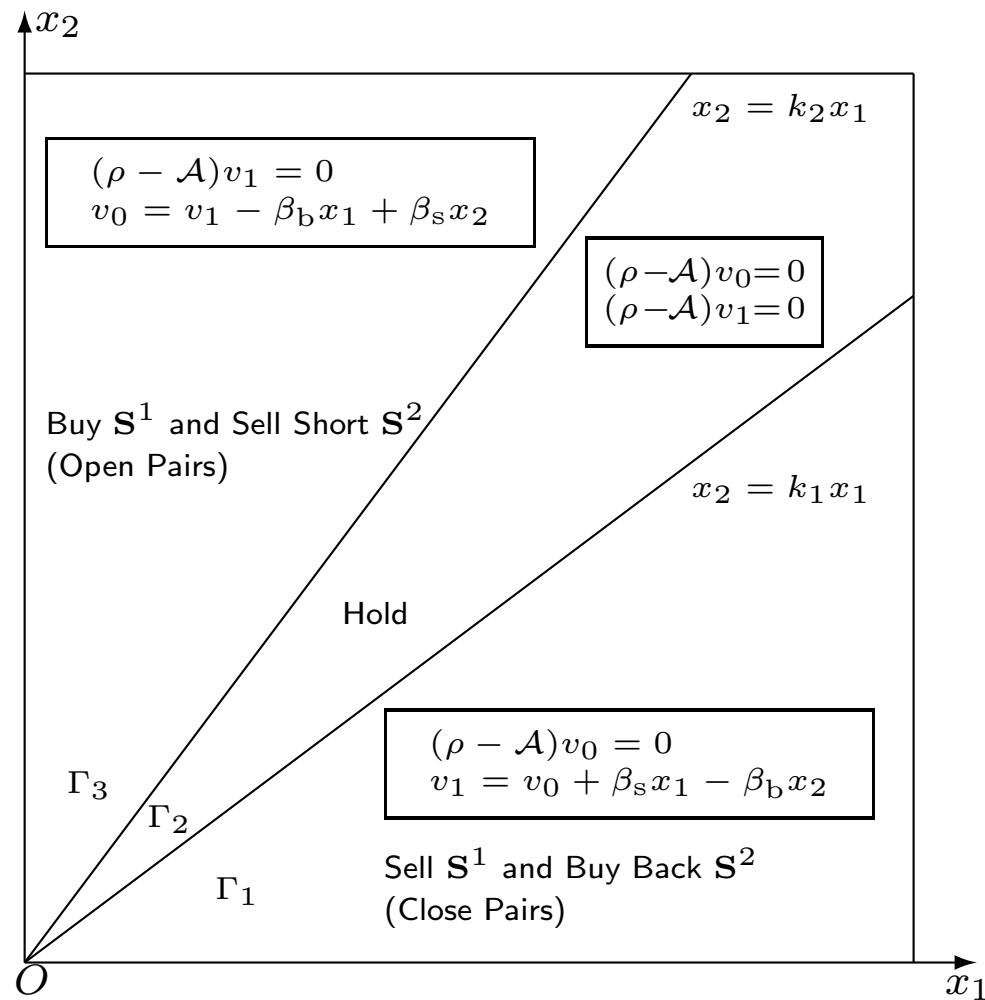
$$\Gamma_1 = \{(x_1, x_2) \in P : x_2 \leq k_1 x_1\},$$

$$\Gamma_2 = \{(x_1, x_2) \in P : k_1 x_1 < x_2 < k_2 x_1\},$$

$$\Gamma_3 = \{(x_1, x_2) \in P : x_2 \geq k_2 x_1\}.$$

Here  $0 < k_1 < k_2$  are slopes to be determined.

# Regions for the Variational Inequalities



## Putting Together the Pieces

---

Recall the boundedness of the value functions and  $\delta_1 > 1$ .  
This implies on  $\Gamma_1$ , the coefficient of  $x_1^{\delta_1} x_2^{1-\delta_1}$  should be 0.  
Therefore,  $v_0 = C_0 x_1^{-\delta_2} x_2^{1+\delta_2}$  for some  $C_0$ .

Likewise, on  $\Gamma_3$ , the coefficient of  $x_1^{-\delta_2} x_2^{1+\delta_2}$  must be zero because  $\delta_2 > 0$ .  
On this region,  $v_1 = C_1 x_1^{\delta_1} x_2^{1-\delta_1}$  for some  $C_1$ .

Finally, these functions are extended to  $\Gamma_2$  and are given by  
 $v_0 = C_0 x_1^{-\delta_2} x_2^{1+\delta_2}$  and  $v_1 = C_1 x_1^{\delta_1} x_2^{1-\delta_1}$ .



# Putting Together the Pieces

---

Therefore,

$$\begin{aligned}\Gamma_1 : \quad v_0 &= C_0 x_1^{-\delta_2} x_2^{1+\delta_2}, & v_1 &= C_0 x_1^{-\delta_2} x_2^{1+\delta_2} + \beta_s x_1 - \beta_b x_2; \\ \Gamma_2 : \quad v_0 &= C_0 x_1^{-\delta_2} x_2^{1+\delta_2}, & v_1 &= C_1 x_1^{\delta_1} x_2^{1-\delta_1}; \\ \Gamma_3 : \quad v_0 &= C_1 x_1^{\delta_1} x_2^{1-\delta_1} - \beta_b x_1 + \beta_s x_2, & v_1 &= C_1 x_1^{\delta_1} x_2^{1-\delta_1}.\end{aligned}$$

## Smooth-Fit Conditions ( $C^1$ conditions).

---

We develop smooth-fit conditions to determine the values for  $k_1$ ,  $k_2$ ,  $C_0$ , and  $C_1$ .

In particular, we are to find  $C^1$  solutions on the entire region  $\{(x_1, x_2) > 0\}$ . Necessarily, the continuous differentiability of  $v_1$  along the line  $x_2 = k_1 x_1$  implies

$$\begin{aligned} C_1 x_1^{\delta_1} x_2^{1-\delta_1} &= C_0 x_1^{-\delta_2} x_2^{1+\delta_2} + \beta_s x_1 - \beta_b x_2, \\ \nabla C_1 x_1^{\delta_1} x_2^{1-\delta_1} &= \nabla (C_0 x_1^{-\delta_2} x_2^{1+\delta_2} + \beta_s x_1 - \beta_b x_2). \end{aligned}$$

This leads to three equations in terms of  $k_1$ ,  $C_0$ , and  $C_1$ . One of these three equations is redundant. Only keep the first two and write

$$\begin{pmatrix} k_1^{1+\delta_2} & -k_1^{1-\delta_1} \\ -\delta_2 k_1^{1+\delta_2} & -\delta_1 k_1^{1-\delta_1} \end{pmatrix} \begin{pmatrix} C_0 \\ C_1 \end{pmatrix} = \begin{pmatrix} \beta_b k_1 - \beta_s \\ -\beta_s \end{pmatrix}.$$

## Smooth-Fit Conditions

---

Similarly, along the line  $x_2 = k_2 x_1$  we have

$$\begin{pmatrix} k_2^{1+\delta_2} & -k_2^{1-\delta_1} \\ -\delta_2 k_2^{1+\delta_2} & -\delta_1 k_2^{1-\delta_1} \end{pmatrix} \begin{pmatrix} C_0 \\ C_1 \end{pmatrix} = \begin{pmatrix} \beta_s k_2 - \beta_b \\ -\beta_b \end{pmatrix}.$$

Write  $C_0$  and  $C_1$  in terms of  $k_1$  and  $k_2$  to obtain:

$$\begin{pmatrix} C_0 \\ C_1 \end{pmatrix} = \frac{1}{\delta_1 + \delta_2} \begin{pmatrix} \beta_s(1 - \delta_1)k_1^{-\delta_2-1} + \beta_b\delta_1 k_1^{-\delta_2} \\ \beta_s(1 + \delta_2)k_1^{\delta_1-1} - \beta_b\delta_2 k_1^{-\delta_1} \end{pmatrix},$$

$$\begin{pmatrix} C_0 \\ C_1 \end{pmatrix} = \frac{1}{\delta_1 + \delta_2} \begin{pmatrix} \beta_b(1 - \delta_1)k_2^{-\delta_2-1} + \beta_s\delta_1 k_2^{-\delta_2} \\ \beta_b(1 + \delta_2)k_2^{\delta_1-1} - \beta_s\delta_2 k_2^{\delta_1} \end{pmatrix}.$$

Eliminating  $C_0$  and  $C_1$ , we obtain two equations for  $k_1$  and  $k_2$ :

$$\begin{pmatrix} \beta_s(1 - \delta_1)k_1^{-\delta_2-1} + \beta_b\delta_1 k_1^{-\delta_2} \\ \beta_s(1 + \delta_2)k_1^{\delta_1-1} - \beta_b\delta_2 k_1^{-\delta_1} \end{pmatrix} = \begin{pmatrix} \beta_b(1 - \delta_1)k_2^{-\delta_2-1} + \beta_s\delta_1 k_2^{-\delta_2} \\ \beta_b(1 + \delta_2)k_2^{\delta_1-1} - \beta_s\delta_2 k_2^{\delta_1} \end{pmatrix}.$$

# Smooth-Fit Conditions

Let  $r = k_2/k_1$ . Then we can show  $r$  is a zero of

$$f(r) = (\delta_1 - 1)\delta_2(\beta_b - \beta_s r^{\delta_1})(\beta_s - \beta_b r^{-\delta_2 - 1}) - \delta_1(1 + \delta_2)(\beta_s - \beta_b r^{\delta_1 - 1})(\beta_b - \beta_s r^{-\delta_2}).$$

We can also show that  $f(1) > 0$  and  $f(\infty) = -\infty$ . There exists  $r_0 > 1$  so that  $f(r_0) = 0$ . Therefore,

$$\begin{cases} k_1 = \frac{(1 + \delta_2)(\beta_s - \beta_b r_0^{\delta_1 - 1})}{\delta_2(\beta_b - \beta_s r_0^{\delta_1})} = \frac{(\delta_1 - 1)(\beta_s - \beta_b r_0^{-\delta_2 - 1})}{\delta_1(\beta_b - \beta_s r_0^{-\delta_2})}, \\ k_2 = \frac{(1 + \delta_2)(\beta_s r_0 - \beta_b r_0^{\delta_1})}{\delta_2(\beta_b - \beta_s r_0^{\delta_1})} = \frac{(\delta_1 - 1)(\beta_s r_0 - \beta_b r_0^{-\delta_2})}{\delta_1(\beta_b - \beta_s r_0^{-\delta_2})}, \end{cases}$$
$$\begin{cases} C_0 = \frac{(1 - \delta_1)\beta_s k_1^{-\delta_2 - 1} + \delta_1\beta_b k_1^{-\delta_2}}{\delta_1 + \delta_2} = \frac{(1 - \delta_1)\beta_b k_2^{-\delta_2 - 1} + \delta_1\beta_s k_2^{-\delta_2}}{\delta_1 + \delta_2}, \\ C_1 = \frac{(1 + \delta_2)\beta_s k_1^{\delta_1 - 1} - \delta_2\beta_b k_1^{\delta_1}}{\delta_1 + \delta_2} = \frac{(\delta_2 + 1)\beta_b k_2^{\delta_1 - 1} - \delta_2\beta_s k_2^{\delta_1}}{\delta_1 + \delta_2}. \end{cases}$$

# Additional Inequalities

---

**Lemma.** We have

$$k_1 < \frac{\beta_s(\rho - \mu_1)}{\beta_b(\rho - \mu_2)}, \quad k_2 > \frac{\beta_b(\rho - \mu_1)}{\beta_s(\rho - \mu_2)}.$$

Also, all inequalities in the HJB equations are satisfied: That is

$$\begin{aligned} (\rho - \mathcal{A})v_0(x_1, x_2) &\geq 0, \\ (\rho - \mathcal{A})v_1(x_1, x_2) &\geq 0, \\ -\beta_b x_1 + \beta_s x_2 &\leq v_0(x_1, x_2) - v_1(x_1, x_2) \leq -\beta_s x_1 + \beta_b x_2. \end{aligned}$$

# Solutions of the HJB Equations

**Theorem.** The solutions of the HJB equations are given by

$$\begin{aligned} v_0(x_1, x_2) &= \begin{cases} \left( \frac{\beta_s(1 - \delta_1)k_1^{-\delta_2-1} + \delta_1\beta_b k_1^{-\delta_2}}{\delta_1 + \delta_2} \right) x_1^{-\delta_2} x_2^{1+\delta_2}, & \text{if } (x_1, x_2) \in \Gamma_1 \cup \Gamma_2, \\ \left( \frac{\beta_s(1 + \delta_2)k_1^{\delta_1-1} - \delta_2\beta_b k_1^{\delta_1}}{\delta_1 + \delta_2} \right) x_1^{\delta_1} x_2^{1-\delta_1} + \beta_s x_2 - \beta_b x_1, & \text{if } (x_1, x_2) \in \Gamma_3, \end{cases} \\ v_1(x_1, x_2) &= \begin{cases} \left( \frac{\beta_s(1 - \delta_1)k_1^{-\delta_2-1} + \delta_1\beta_b k_1^{-\delta_2}}{\delta_1 + \delta_2} \right) x_1^{-\delta_2} x_2^{1+\delta_2} + \beta_s x_1 - \beta_b x_2, & \text{if } (x_1, x_2) \in \Gamma_1, \\ \left( \frac{\beta_s(1 + \delta_2)k_1^{\delta_1-1} - \delta_2\beta_b k_1^{\delta_1}}{\delta_1 + \delta_2} \right) x_1^{\delta_1} x_2^{1-\delta_1}, & \text{if } (x_1, x_2) \in \Gamma_2 \cup \Gamma_3. \end{cases} \end{aligned}$$

# A Verification Theorem

**Theorem.** (a) We have  $v_i(x) = V_i(x)$ ,  $i = 0, 1$ .

(b) Moreover, if initially  $i = 0$ , let  $\Lambda_0^* = (\tau_1^*, \tau_2^*, \tau_3^*, \dots)$  such that  $\tau_1^* = \inf\{t \geq 0 : (X_t^1, X_t^2) \in \Gamma_3\}$ ,  $\tau_2^* = \inf\{t \geq \tau_1^* : (X_t^1, X_t^2) \in \Gamma_1\}$ ,  $\tau_3^* = \inf\{t \geq \tau_2^* : (X_t^1, X_t^2) \in \Gamma_3\}$ , and so on. Similarly, if initially  $i = 1$ , let  $\Lambda_1^* = (\tau_0^*, \tau_1^*, \tau_2^*, \dots)$  such that  $\tau_0^* = \inf\{t \geq 0 : (X_t^1, X_t^2) \in \Gamma_1\}$ ,  $\tau_1^* = \inf\{t \geq \tau_0^* : (X_t^1, X_t^2) \in \Gamma_3\}$ ,  $\tau_2^* = \inf\{t \geq \tau_1^* : (X_t^1, X_t^2) \in \Gamma_1\}$ , and so on. Then  $\Lambda_0^*$  and  $\Lambda_1^*$  are optimal.

*Proof.* Key steps:

- (1)  $v_0(x_1, x_2) \geq 0$ .
- (2)  $v_0(x_1, x_2)$  and  $v_1(x_1, x_2)$  satisfy linear growth conditions.
- (3)  $\tau_n^* \rightarrow \infty$ , a.s.
- (4)  $v_i(x_1, x_2) \geq J_i(x_1, x_2, \Lambda_i)$ ,  $i = 0, 1$ .
- (5)  $v_i(x_1, x_2) = J_i(x_1, x_2, \Lambda_i^*)$ ,  $i = 0, 1$ .

## A Numerical Example

---

Use the parameters of the TGT-WMT example:

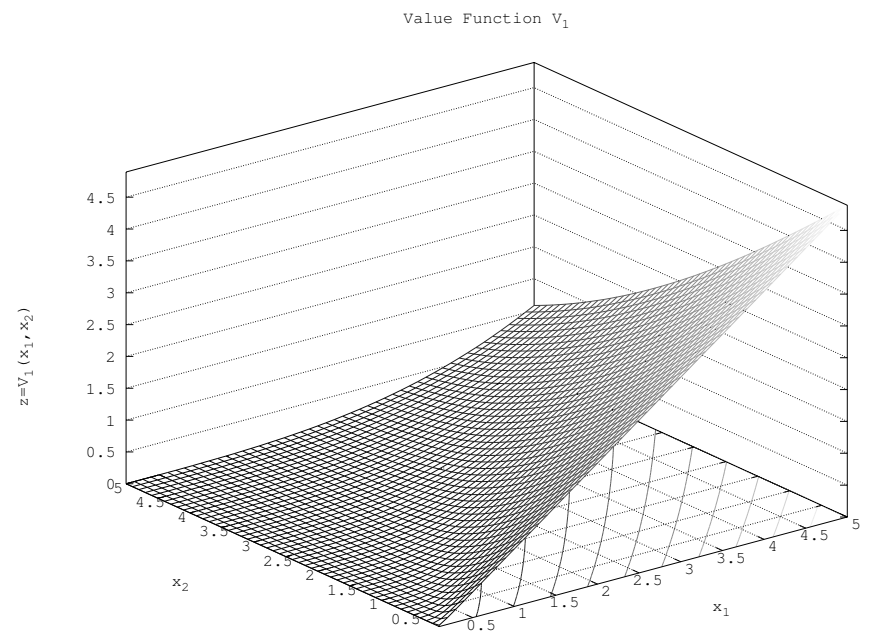
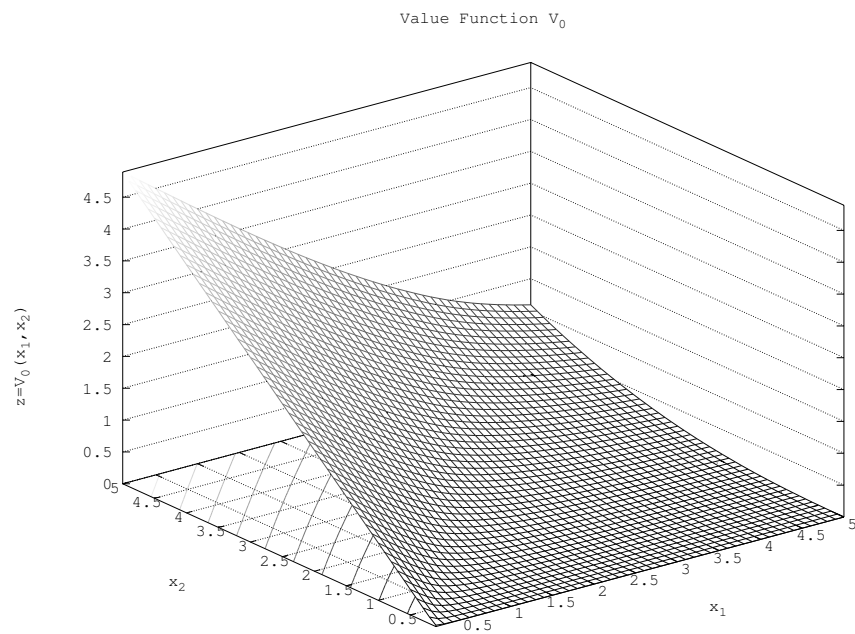
$$\mu_1 = 0.2059, \mu_2 = 0.2459, \sigma_{11} = 0.3112, \sigma_{12} = 0.0729, \sigma_{21} = 0.0729, \sigma_{22} = 0.2943.$$

Take  $K = 0.001$  and  $\rho = 0.5$ . We obtain

$$k_1 = 1.03905, \quad k_2 = 1.28219.$$



# Value Functions



## Dependence of $(k_1, k_2)$ on Parameter $\mu_1$

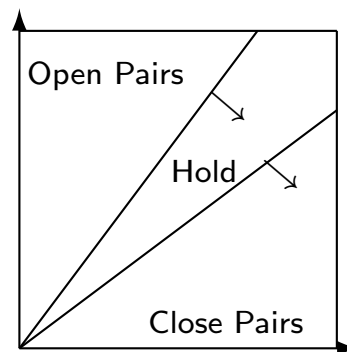
We vary one of the parameters at a time and examine the dependence of  $(k_1, k_2)$  on these parameters.

First we consider how  $(k_1, k_2)$  changes with  $\mu_1$ . A larger  $\mu_1$  implies greater potential of growth in  $S^1$ .

It can be seen in Table 1 that both  $k_1$  and  $k_2$  decrease in  $\mu_1$  leading to more buying opportunities.

$\mu_1$	0.1059	0.1559	0.2059	0.2559	0.3059
$k_1$	1.38860	1.21356	1.03905	0.86272	0.68532
$k_2$	1.70104	1.49268	1.28219	1.07150	0.86008

Table 1.  $(k_1, k_2)$  with varying  $\mu_1$ .

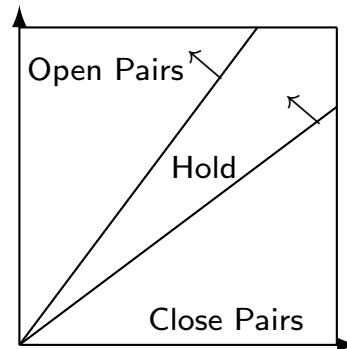


## Dependence of $(k_1, k_2)$ on Parameter $\mu_2$

Next, we vary  $\mu_2$ . It is clear in this case that the pair  $(k_1, k_2)$  increase in  $\mu_2$ . This is because larger  $\mu_2$  means bigger growth potential in  $\mathbf{S}^2$  which discourages establishing pairs position  $\mathbf{Z}$  and encourages its early exit.

$\mu_2$	0.1459	0.1959	0.2459	0.2959	0.3459
$k_1$	0.75424	0.87372	1.03905	1.28131	1.67831
$k_2$	0.92168	1.07205	1.28219	1.59780	2.11803

Table 2.  $(k_1, k_2)$  with varying  $\mu_2$ .



## Dependence of $(k_1, k_2)$ on Parameters $\sigma_{11}$ and $\sigma_{22}$

In Tables 3 and 4, we vary the volatility  $\sigma_{11}$  and  $\sigma_{22}$ .

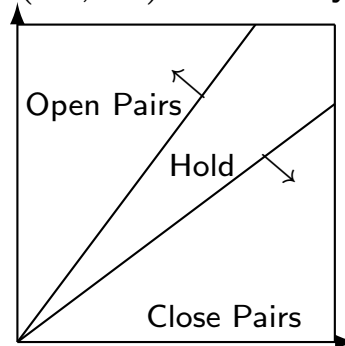
Larger volatility leads higher risk, which translates to smaller buying zone  $\Gamma_3$ .

On the other hand, larger volatility gives more room for the price to move.

This leads to smaller selling zone  $\Gamma_1$ .

$\sigma_{11}$	0.2112	0.2612	0.3112	0.3612	0.4112
$k_1$	1.05320	1.04598	1.03905	1.02997	1.02008
$k_2$	1.26384	1.27295	1.28219	1.29364	1.30417
$\sigma_{22}$	0.1943	0.2443	0.2943	0.3443	0.3943
$k_1$	1.05147	1.04511	1.03905	1.03224	1.02469
$k_2$	1.26597	1.27399	1.28219	1.29133	1.30136

Tables 3 and 4.  $(k_1, k_2)$  with varying  $\sigma_{11}$  and  $\sigma_{22}$ .

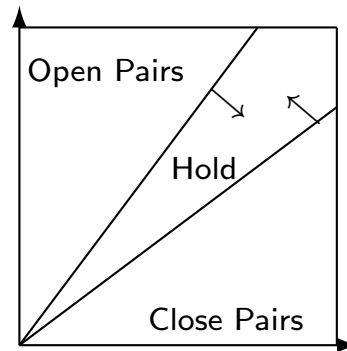


## Dependence of $(k_1, k_2)$ on Parameter $\sigma_{12}$

Next, we vary  $\sigma_{12}$  which equals  $\sigma_{21}$ . Note that this parameter dictates the correlation between  $X_t^1$  and  $X_t^2$ . Larger  $\sigma_{12}$  leads to greater correlation, which encourages more buying opportunities (larger  $\Gamma_3$ ) and more selling as well (larger  $\Gamma_1$ ).

$\sigma_{12}(= \sigma_{21})$	-0.0271	0.0229	0.0729	0.1229	0.1729
$k_1$	1.00965	1.02318	1.03905	1.05546	1.07276
$k_2$	1.32062	1.30251	1.28219	1.26127	1.23904

Table 5.  $(k_1, k_2)$  with varying  $\sigma_{12}(= \sigma_{21})$ .

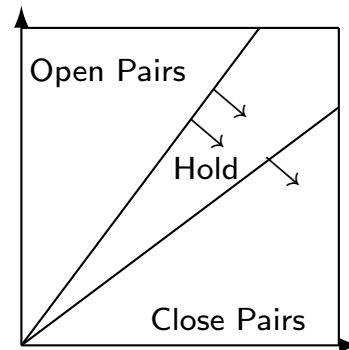


## Dependence of $(k_1, k_2)$ on Parameter $\rho$

Next, we vary the discount rate  $\rho$ . Larger  $\rho$  encourages quicker profits, which leads to more buying and shorter holding. This is confirmed in Table 6. It shows that larger  $\rho$  leads to a smaller  $k_2$  and smaller  $(k_2 - k_1)$ .

$\rho$	0.40	0.45	0.50	0.55	0.60
$k_1$	1.10935	1.06547	1.03905	1.02291	1.00997
$k_2$	1.41886	1.33396	1.28219	1.24591	1.22105
$k_2 - k_1$	0.30951	0.26849	0.24314	0.22300	0.21108

Table 6.  $(k_1, k_2)$  with varying  $\rho$ .

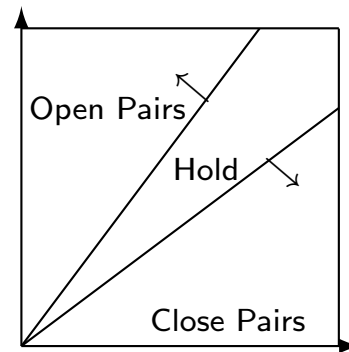


## Dependence of $(k_1, k_2)$ on Parameter $K$

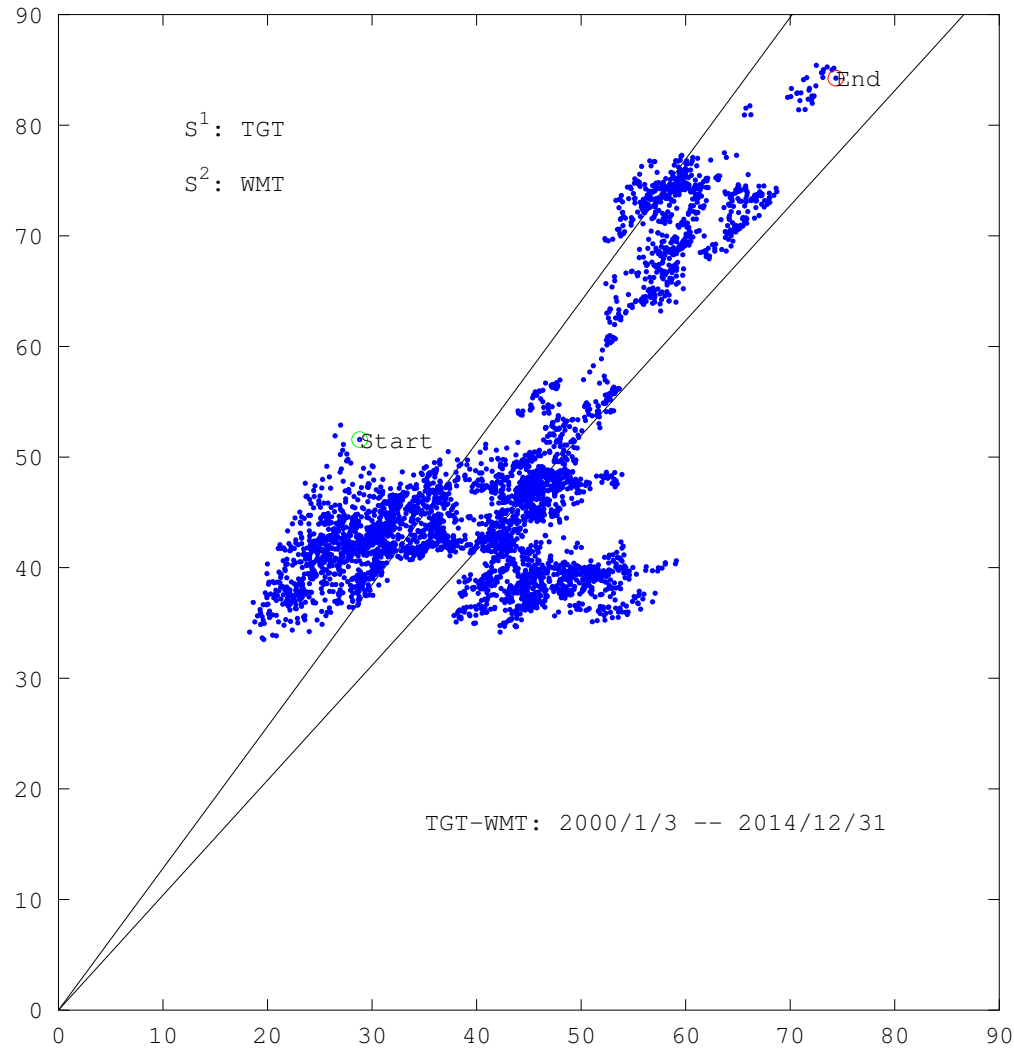
Finally, we examine the dependence on the transaction percentage  $K$ . Clearly, a larger  $K$  discourages trading transactions. This results smaller buying zone  $\Gamma_3$  and smaller selling zone  $\Gamma_1$ .

$K$	0.0001	0.0005	0.001	0.002	0.003
$k_1$	1.07951	1.06318	1.03905	1.00787	0.98627
$k_2$	1.23819	1.25562	1.28219	1.31728	1.34231

Table 7.  $(k_1, k_2)$  with varying  $K$ .



# Backtesting (TGT-WMT): Daily Closing Price Distribution





## Backtesting (TGT–WMT)

---

Using the parameters obtained earlier based on the historical prices from 1985 to 1999, we obtain  $(k_1, k_2) = (1.03905, 1.28219)$ .

A pairs trading (long  $S^1$  and short  $S^2$ ) is triggered when  $(X_t^1, X_t^2)$  enters  $\Gamma_3$ .

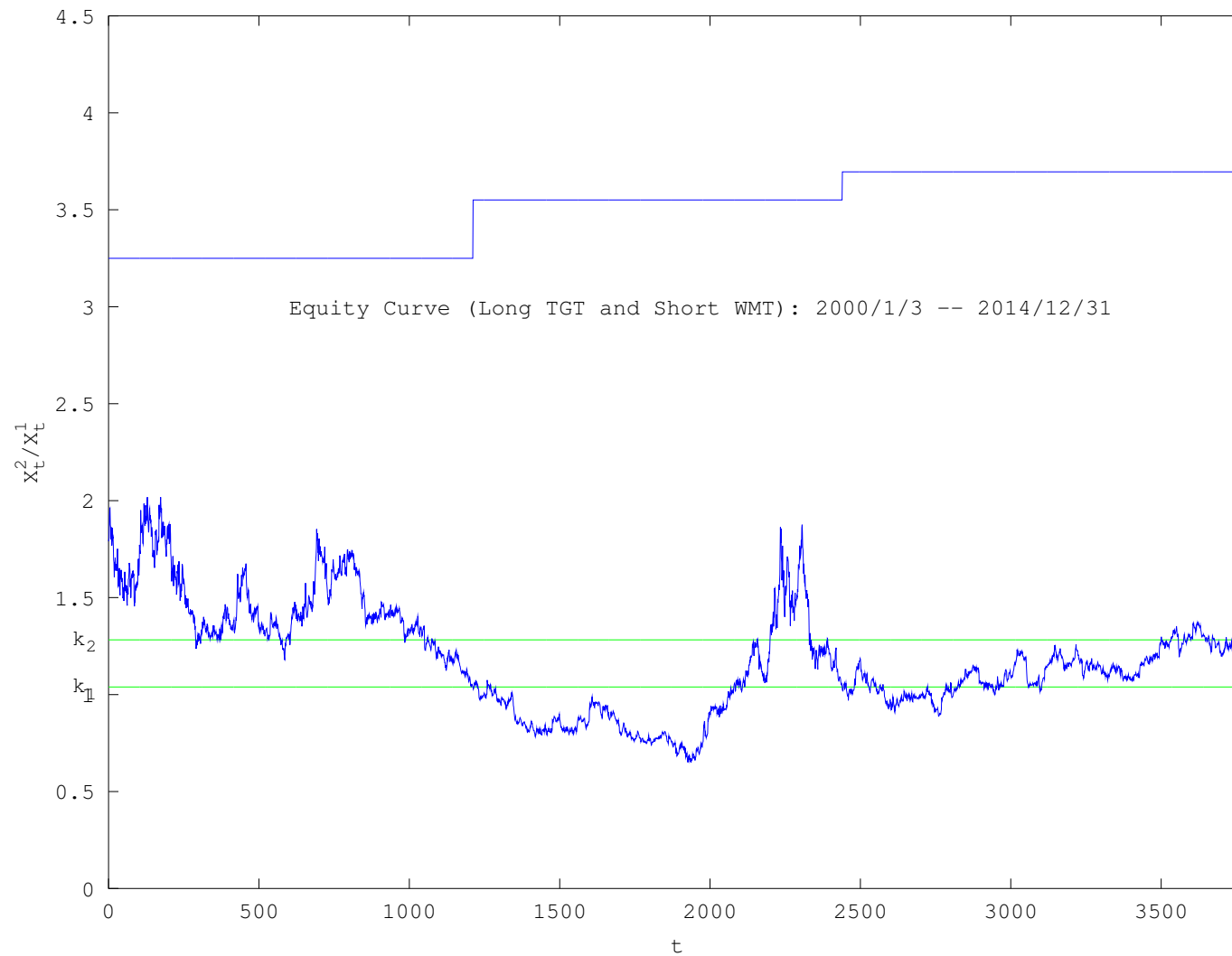
The position is closed when  $(X_t^1, X_t^2)$  enters  $\Gamma_1$ .

Initially, allocate trading capital \$100K. When the first long signal is triggered, buy \$50K TGT stocks and short the same amount of WMT.

Such half-and-half capital allocation between long and short applies to all trades.

Each pairs transaction is charged \$5 commission.

# Backtesting (TGT-WMT): $S^1 = \text{TGT}$ , $S^2 = \text{WMT}$



## Backtesting (TGT–WMT): $S^1 = \text{TGT}$ , $S^2 = \text{WMT}$

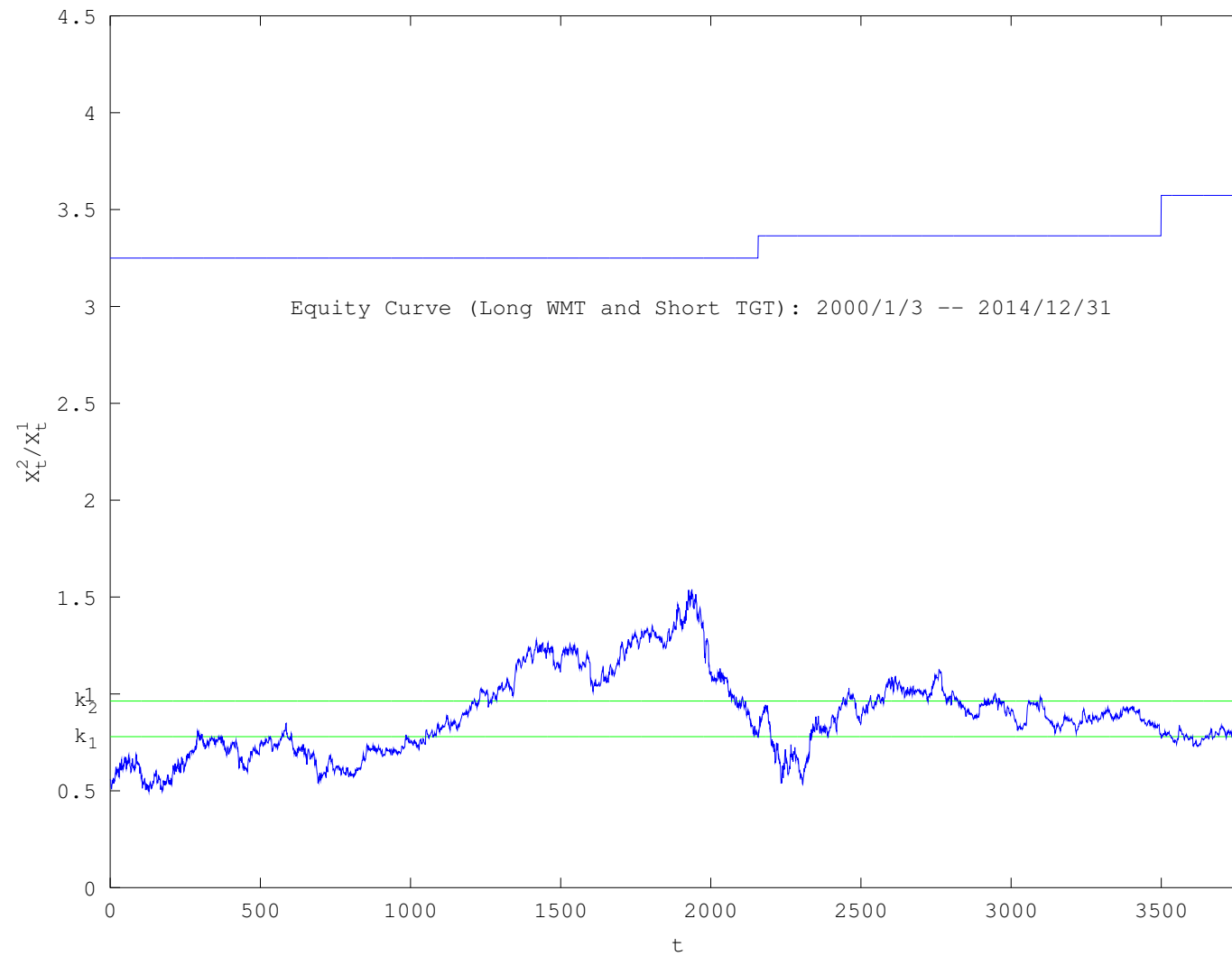
---

There were three trades with end balance \$155.914K.

We can also switch the roles of  $S^1$  and  $S^2$ , i.e., to long WMT and short TGT by taking  $S^1 = \text{WMT}$  and  $S^2 = \text{TGT}$ .

In this case, the new  $(\tilde{k}_1, \tilde{k}_2) = (1/k_2, 1/k_1) = (1/1.28219, 1/1.03905)$ .

# Backtesting (TGT-WMT): $S^1=WMT$ , $S^2=TGT$



## Backtesting (TGT–WMT): $S^1=WMT$ , $S^2=TGT$

---

Such trade leads to the end balance \$132.340K.

Note that both types of trades have no overlap, i.e., they do not compete for the same capital. The grand total profit is \$88254 which is a 88.25% gain.

The main advantage of pairs trading is its risk neutral nature, i.e., it can be profitable regardless general market conditions.

## Extension: Pairs trading with cutting losses

---

To limit downside risk of a pairs position, we impose a hard cut loss level. Any existing position will be automatically closed upon entering the cut loss region.

Trading with cutting losses is important in practice to limit risk exposure due to unexpected events.

A stop-loss limit is often enforced as part of money management.

It can also be associated with a margin call due to substantial losses.

In control theory, this is associated with a hard state constraint which is difficult to deal with.

---

In particular, we impose state constraint and require  $S_t^2/S_t^1 \leq M$ .

Here  $M$  is a given constant representing a stop-loss level to account for unforeseeable event in the marketplace.

Let  $\tau_M = \{t : S_t^2/S_t^1 \geq M\}$ .

Then,  $\tau_n^S \leq \tau_M$  and  $\tau_n^B \leq \tau_M$ , for all  $n$ .

Let  $\Lambda_0 = (\tau_1^B, \tau_1^S, \tau_2^B, \dots)$  and  $\Lambda_1 = (\tau_0^S, \tau_1^B, \tau_1^S, \dots)$ .

Goal: To find  $\Lambda_0$  and  $\Lambda_1$  so as to maximize

$$\left\{ \begin{array}{l} J_0(x_1, x_2, \Lambda_0) = E \left\{ [e^{-\rho\tau_1^S} (\beta_s S_{\tau_1^S}^1 - \beta_b S_{\tau_1^S}^2) - e^{-\rho\tau_1^B} (\beta_b S_{\tau_1^B}^1 - \beta_s S_{\tau_1^B}^2)] I_{\{\tau_1^B < \tau_M\}} \right. \\ \quad \left. + [e^{-\rho\tau_2^S} (\beta_s S_{\tau_2^S}^1 - \beta_b S_{\tau_2^S}^2) - e^{-\rho\tau_2^B} (\beta_b S_{\tau_2^B}^1 - \beta_s S_{\tau_2^B}^2)] I_{\{\tau_2^B < \tau_M\}} \right] + \dots \right\}, \\ J_1(x_1, x_2, \Lambda_0) = E \left\{ e^{-\rho\tau_0^S} (\beta_s S_{\tau_0^S}^1 - \beta_b S_{\tau_0^S}^2) \right. \\ \quad \left. + [e^{-\rho\tau_1^S} (\beta_s S_{\tau_1^S}^1 - \beta_b S_{\tau_1^S}^2) - e^{-\rho\tau_1^B} (\beta_b S_{\tau_1^B}^1 - \beta_s S_{\tau_1^B}^2)] I_{\{\tau_1^B < \tau_M\}} \right] + \dots \right\}. \end{array} \right.$$



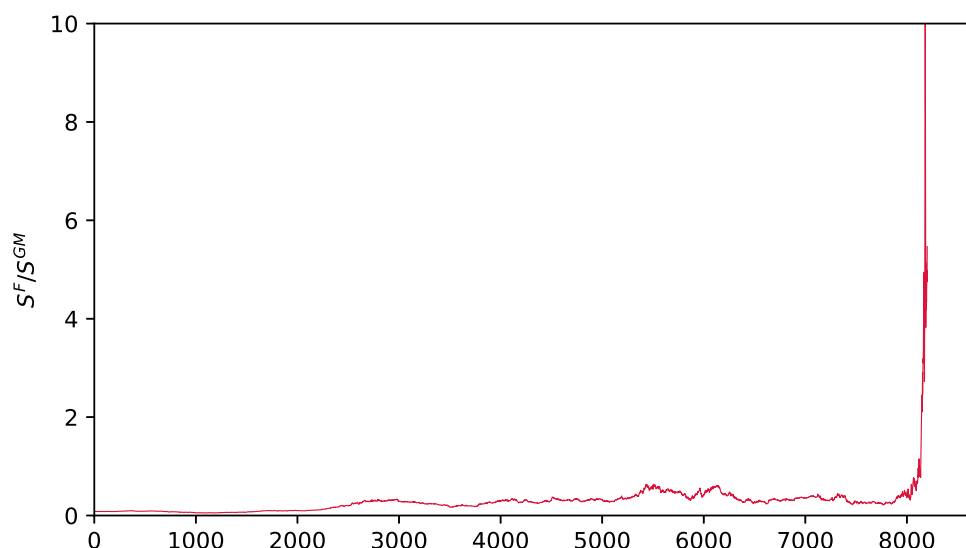
## F vs GM: Daily closing ratio of F/GM from 1977 to 2009

The stock prices of Ford (F) and GM (GM) are highly correlated historically. (good candidates for pairs trading).

The ratio remains 'normal' until it is not (when it approaches the past subprime crisis).

This would trigger a pairs position (long GM and short F).

It spikes prior to GM's chapter 11 filing on 6/1/2009 causing heavy losses to any F/GM pair positions. Necessary to set up stop loss limits.



# HJB equations

---

For  $x_1 > 0$  and  $0 < x_2 < Mx_1$ ,

$$\begin{cases} \min\{\rho v_0(x_1, x_2) - \mathcal{A}v_0(x_1, x_2), v_0(x_1, x_2) - v_1(x_1, x_2) + \beta_b x_1 - \beta_s x_2\} = 0, \\ \min\{\rho v_1(x_1, x_2) - \mathcal{A}v_1(x_1, x_2), v_1(x_1, x_2) - v_0(x_1, x_2) - \beta_s x_1 + \beta_b x_2\} = 0, \end{cases}$$

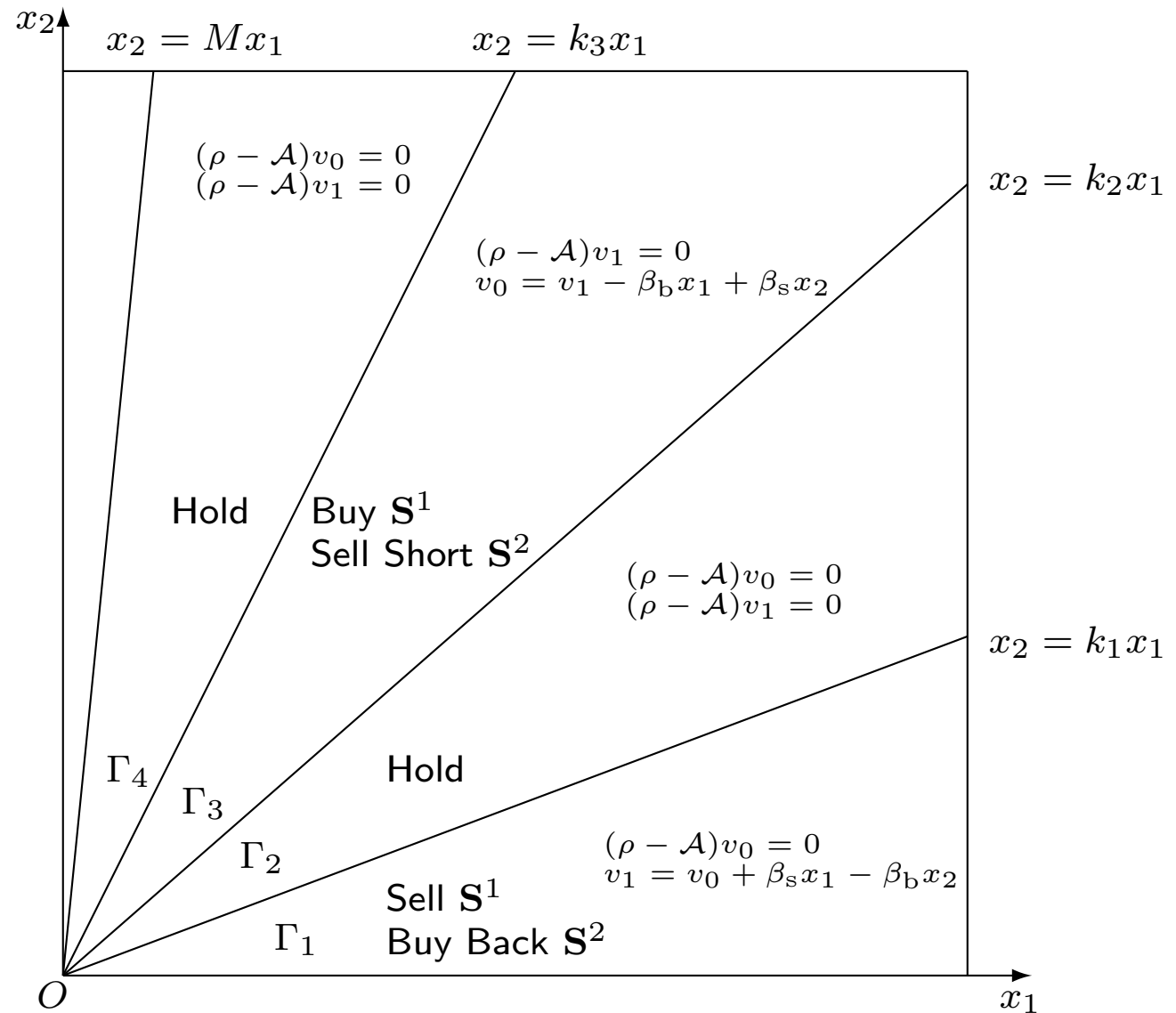
Boundary conditions:  $v_0(x_1, Mx_1) = 0$  and  $v_1(x_1, Mx_1) = \beta_s x_1 - \beta_b Mx_1$ .

Let  $y = \frac{x_2}{x_1}$  and  $v_i(x_1, x_2) = x_1 w_i(\frac{x_2}{x_1})$  for some  $w_i(\cdot)$  for  $i = 0, 1$ .

$$\mathcal{L}(w_i(y)) = \lambda y^2 w_i''(y) + (\mu_2 - \mu_1) y w_i'(y) + \mu_1 w_i(y).$$

$$\begin{cases} \min\{\rho w_0(y) - \mathcal{L}w_0(y), w_0(y) - w_1(y) + \beta_b - \beta_s y\} = 0, \\ \min\{\rho w_1(y) - \mathcal{L}w_1(y), w_1(y) - w_0(y) - \beta_s + \beta_b y\} = 0, \\ w_0(M) = 0, \quad w_1(M) = \beta_s - \beta_b M. \end{cases}$$

# Regions $\Gamma_1, \Gamma_2, \Gamma_3$ , and $\Gamma_4$



# Assumptions

(A1)  $\rho > \mu_1$  and  $\rho > \mu_2$ .

(A2) There is  $k_3$  in  $(k_2, M)$  such that  $h_1(k_3) = 0$ ,  
where

$$h_1(x) = \frac{M^{\delta_1} \beta_s (x(1 - \delta_2) + \beta \delta_2)}{x^{\delta_1}} + \frac{M^{\delta_2} \beta_s (x(\delta_1 - 1) - \beta \delta_1)}{x^{\delta_2}} + \beta_s (1 - M\beta)(\delta_1 - \delta_2),$$

$$\delta_1 = \frac{1}{2} \left( 1 + \frac{\mu_1 - \mu_2}{\lambda} + \sqrt{\left( 1 + \frac{\mu_1 - \mu_2}{\lambda} \right)^2 + \frac{4\rho - 4\mu_1}{\lambda}} \right),$$
$$\delta_2 = \frac{1}{2} \left( 1 + \frac{\mu_1 - \mu_2}{\lambda} - \sqrt{\left( 1 + \frac{\mu_1 - \mu_2}{\lambda} \right)^2 + \frac{4\rho - 4\mu_1}{\lambda}} \right).$$

(Sufficient:  $h_1(k_2) > 0$ ).

(A3)  $h'_2(M) < 0$  or  $h''_2(M) < 0$ ,

where

$$h_2(y) = w_1 - w_0 + \beta_b y - \beta_s.$$

(Sufficient:  $\mu_1 \geq \mu_2$ ).

# Theorem

---

Assume (A1), (A2), and (A3). Then the following functions  $v_i(x_1, x_2) = x_1 w_i(x_2/x_1)$ ,  $i = 0, 1$ , satisfy the HJB equations:

$$w_0(y) = \begin{cases} C_0 y^{\delta_1}, & \text{for } 0 < y \leq k_2, \\ C_1 y^{\delta_1} + C_2 y^{\delta_2} + \beta_s y - \beta_b, & \text{for } k_2 < y \leq k_3, \\ C_3 y^{\delta_1} + C_4 y^{\delta_2}, & \text{for } k_3 < y \leq M; \end{cases}$$
$$w_1(y) = \begin{cases} C_0 y^{\delta_1} + \beta_s - \beta_b y, & \text{for } 0 < y \leq k_1, \\ C_1 y^{\delta_1} + C_2 y^{\delta_2}, & \text{for } k_1 < y \leq M, \end{cases}$$

# Theorem

Assume (A1), (A2), and (A3) and  $v_0(x_1, x_2) \geq 0$ . Then,

$$v_i(x_1, x_2) = x_1 w_i(x_2/x_1) = V_i(x_1, x_2), \quad i = 0, 1.$$

Moreover, if  $i = 0$ , let  $\Lambda_0^* = (\tau_1^{B*}, \tau_1^{S*}, \tau_2^{B*}, \dots) = (\tau_1^{B,0}, \tau_1^{S,0}, \tau_2^{B,0}, \dots) \wedge \tau_M$

where  $\tau_1^{B,0} = \inf\{t \geq 0 : (S_t^1, S_t^2) \in \Gamma_3\}$ ,

$$\tau_1^{S,0} = \inf\{t \geq \tau_1^{B,0} : (S_t^1, S_t^2) \in \Gamma_1\},$$

$$\tau_2^{B,0} = \inf\{t \geq \tau_1^{S,0} : (S_t^1, S_t^2) \in \Gamma_3\}, \dots$$

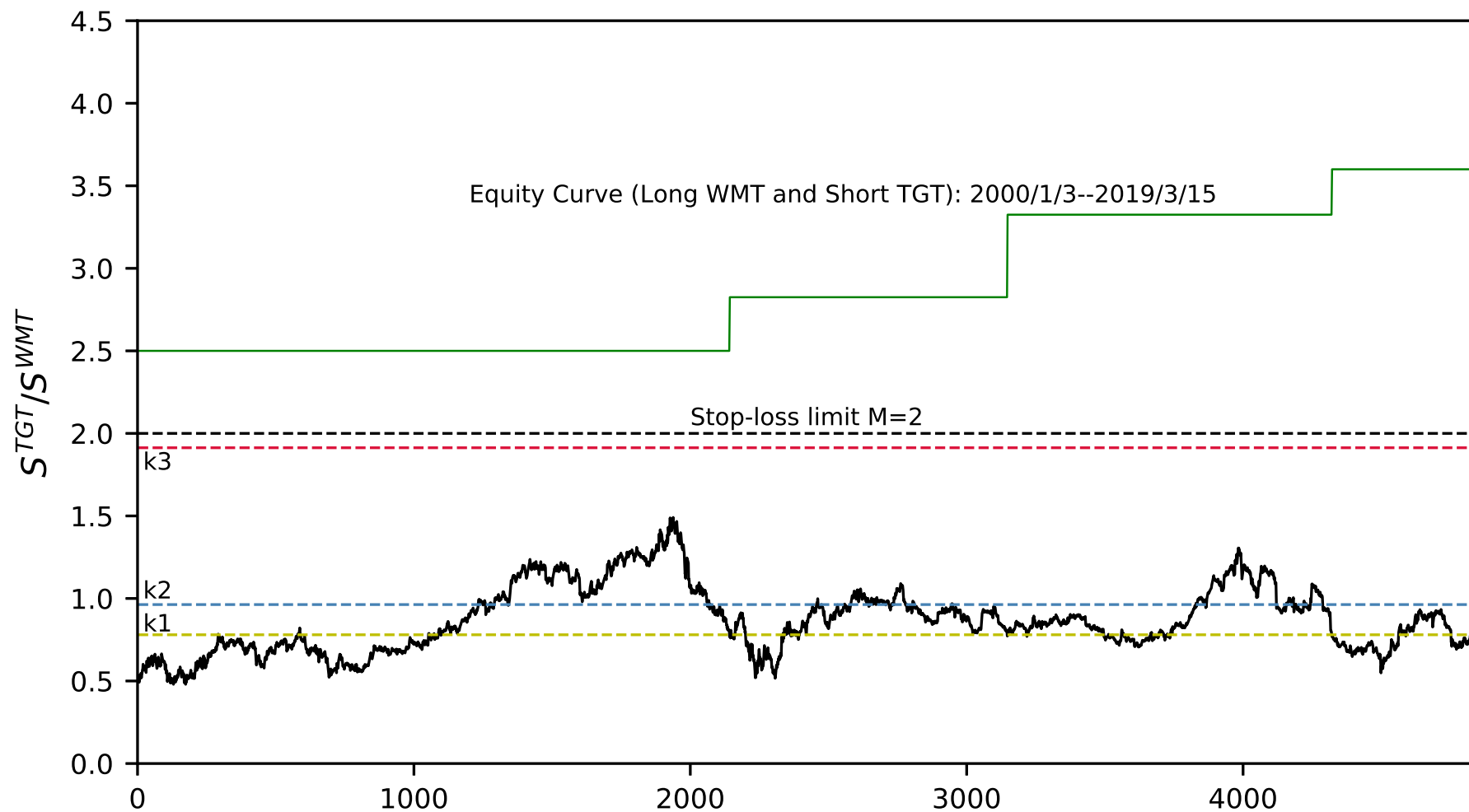
Similarly, if  $i = 1$ , let  $\Lambda_1^* = (\tau_0^{S*}, \tau_1^{B*}, \tau_1^{S*}, \dots) = (\tau_0^{S,0}, \tau_1^{B,0}, \tau_1^{S,0}, \dots) \wedge \tau_M$

where  $\tau_0^{S,0} = \inf\{t \geq 0 : (S_t^1, S_t^2) \in \Gamma_1\}$ ,

$$\tau_1^{B,0} = \inf\{t \geq \tau_0^{S,0} : (S_t^1, S_t^2) \in \Gamma_3\}, \quad \tau_1^{S,0} = \inf\{t \geq \tau_1^{B,0} : (S_t^1, S_t^2) \in \Gamma_1\}, \dots$$

Then  $\Lambda_0^*$  and  $\Lambda_1^*$  are optimal.

$S^1 = \text{WMT}$ ,  $S^2 = \text{TGT}$ : with cutloss  $M = 2$



# Conclusion

---

A pairs trading problem was studied under traditional GBMs.

A **closed-form** solution was obtained.

It would be interesting to study pairs trading under more **realistic models**, e.g., GBM with regime switching with possible partial observation [update (Pairs selling rule under regime switching with full observation: Done; General trading rules with full observation: in progress).]

It would also be interesting to examine how the method works for a **larger selection** of pairs of stocks at possible **different time scales**.

**Some other types of pairs selections:** stock + call (put) option, a stock portfolio + another stock portfolio, a stock portfolio + a stock index future (or ETF), (e.g., long a portfolio with under valued stocks and short an index ETF)...



# References

---

- McDonald and Siegel. 1986. The value of waiting to invest, *The Quarterly Journal of Economics*.
- Hu and Øksendal. 1998. Optimal time to invest when the price processes are geometric Brownian motions, *Finance and Stochastics*.
- Vidyamurthy. 2004. *Pairs Trading: Quantitative Methods and Analysis*, Wiley, Hoboken, NJ.
- Gatev, Goetzmann, and Rouwenhorst. 2006. Pairs trading: Performance of a relative-value arbitrage rule, *Review of Financial Studies*.
- Zhang and Zhang. 2008. Trading a mean-reverting asset: Buy low and sell high, *Automatica*.
- Song and Zhang, 2013. An optimal pairs-trading rule, *Automatica*.
- Tie, Zhang, and Zhang, 2018. An Optimal strategy for pairs-trading under geometric Brownian motions, *Journal of Optimization Theory and Applications*.
- Liu, Wu, and Zhang, (submitted). Pairs-trading under geometric Brownian motions with cutting losses.
- J. Tie and Q. Zhang, (submitted). Pairs trading: An optimal selling rule under a regime switching model.