The Non-Misleading Value of Inferred Correlation: An Introduction to the Cointelation Model

Babak Mahdavi Damghani
e-mail: bmd@cantab.net

Abstract
This paper[1] comes together with growing evidence for power law-type scaling[2] of correlation with time, a concept rooted in the original study by Benoit Mandelbrot on concentration of risk.[3] We complete the cointelation model recently introduced[4] via a statistical test, which uses measured correlation in different time gaps. We also provide an approximation of the expectation in the change in measured correlation via these various time steps and use our findings in order to introduce the concept of inferred correlation and the term structure of correlation. We finally illustrate our findings through the example of the relation between oil and BP, and present a few potential applications in the financial industry.

Keywords
cointelation, inferred correlation, measured correlation, cointegration, correlation term structure

1 Introduction

1.1 Scope

Within the context of, arguably, a mathematician-led financial crisis, Gosset [1] showed that the industry needed to take a step back and reflect on some of its core assumptions. With this in mind Damghani *et al.* [2] explained how, within the framework of the financial industry, when representing relationships between assets, correlation is typically used in lieu of perhaps sometimes a more realistic model that they named cointelation. Cointelation is a portmanteau neologism in finance, designed to signify a hybrid method between the cointegration[5] and correlation models. While this cointelation model was not the main point the authors wanted to introduce, it however provided an original model for which a methodology for the estimation of the parameters was addressed; but no test was suggested in order to determine whether two stochastic processes were "cointelated" to begin with. The objectives of the paper are first, to give the specifications for this cointelation test and second, to try to reconcile this new idea with a market observable phenomenon – which is that the correlation on different time gaps increases as the time between returns increases for certain pairs. Last, we introduce the concept of inferred correlation which will be defined as its closest immediately translatable industry cousin for the cointelated pairs.

1.2 Structure of the paper

In Section 2 we give a reminder of the cointelation model as well as the models it was inherited from. In Section 3 we lay out the intuition and specifications for the cointelation test. In Section 4 we introduce the concept of inferred correlation. In Section 5 we give the specification of the cointelation test. In Section 6 we explain the parameter estimation methodology. We apply the test to the example of oil and BP in Section 7. Finally, in Section 8, we give a few examples of applications in the financial industry.

2 Correlation, cointegration, and cointelation

2.1 Model set-up and review

The naming of the parameters and the model set-up are those taken from "The misleading value of measured correlation" [2]. That is, we set up the probability space $(\Omega, (f)_{(t\geq 0)}, \mathbb{P})$, with $(f)_{(t\geq 0)}$ generated by the $(T+1)$-dimensional Brownian motion and $\mathbb{P}$ the historical probability measure under which the discounted price of the underlier, rS, is not necessarily a martingale. The main objective of this paper not being pricing, we will revert from working in the risk-neutral probability space.

2.2 Leading stochastic process

We first define our leading stochastic process, $S_t (t \geq 0)$, in eqn (1). We have defined this stochastic process to be the "leading" stochastic process because the movements of the other three models we introduce in eqns (2), (4), and (5) are all conditional on the movements of this leading stochastic process. Alternatively, one may want to think of this process as the combined products of macro events which are unpredictable.

$$\frac{dS_t}{S_t} = r dt + \sigma dW_t \tag{1}$$

with $dW_t \sim \mathbb{N}(0, dt)$. Here, the distribution of the errors has been chosen arbitrarily. It obviously does not need to be lognormal. One may want to choose a fat-tailed distribution, especially if one considers the leading indicator to be representing macro events.

2.3 Model: Correlation

The second stochastic process, $S_{r,t}$ $(t \geq 0)$, that we have labeled "the correlation model," is defined by eqn (2). In this equation, dW_t is the same as in eqn (1) and $dW_t^{\perp}$ is an independent Brownian motion. In this example the volatility has been chosen to be the same as in eqn (1) so as to avoid the types of situation Paul Wilmott described [3]. More specifically, situations in which correlation was only relevant at the infinitesimal level and the resulting measured correlation could be misleading with respect to the relative departure of the two relevant underliers.

$$\frac{dS_t}{S_t} = rdt + \sigma dW_t$$
$$\frac{dS_{r,t}}{S_{r,t}} = \sigma(\rho dW_t + \sqrt{1-\rho^2}\, dW_t^{\perp}) \qquad (2)$$

with $dW_t^{\perp} \sim \mathbb{N}(0, dt)$ and ρ representing the correlation coefficient between the two stochastic processes at the most infinitesimal level. In this technical paper, measured correlation will refer to Pearson's correlation coefficient [represented by eqn (3)]:

$$\rho_{xy} = \frac{\sigma_{xy}^2}{\sigma_x \sigma_y} \qquad (3)$$

2.4 Model: Cointegration

We know that two or more time series are cointegrated if they have the same stochastic drift. There are different ways to represent such a relationship. We have chosen here to use one inspired by the Ornstein–Uhlenbeck (OU) model [4],[6] where the stochastic differential equation is described by eqn (4). In this model, θ represents the recall force or the speed at which the "lagging" stochastic process [here generated by eqn (4)] mean reverts around the leading stochastic process [here generated by eqn (1)]. Note that $\theta \in]0,1]$. Indeed, if $\theta = 0$ we are back to the first model and if $\theta > 1$ the stochastic process would "explode." The $S_{g,t}$ in eqn (4) is there to prevent the stochastic process ever going negative.

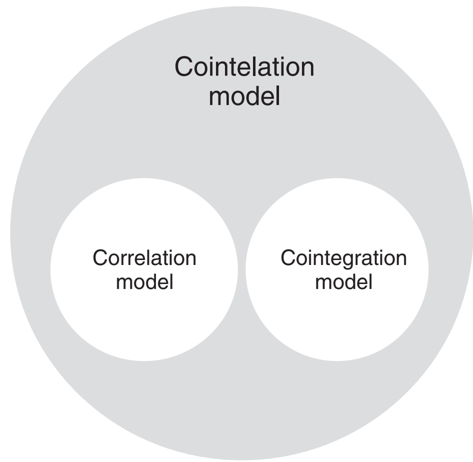
$$\frac{dS_t}{S_t} = rdt + \sigma dW_t$$
$$dS_{g,t} = \theta(S_t - S_{g,t})dt + \sigma S_{g,t} dW_t^{\perp} \qquad (4)$$

One may argue that this model is restrictive in the sense that it assumes mean reversion. This perception has been described and explained as the correlation bias [2]. This model remains as relevant as the correlation model and needs to be taken into account also when it comes to representing the relationship between two assets.

2.5 Model: Cointelation

The cointelation model, as explained in the Introduction, is a hybrid model between the correlation and cointegration models specified in eqns (2) and (4). The idea of the model is that we would like to represent two stochastic processes with the same long-term drift (as is the case for the cointegration model) but also have a model that would appear to be behaving like the correlation model most of the time or at least in the short term. This dual objective is achieved by combining the main properties of the two models. In that sense, the cointelation model could be seen as a generalization (Figure 1) of both the correlation and the cointegration model, and as such it cannot be "restrictive." The long-term drift argument is achieved by using the deterministic part of the differential equation, the same way it is used in the cointegration model: via a recall force fueled by the θ parameter. The second objective, that is a relationship in the short term, is reached when the lagging stochastic process (S_l) is close to its leading indicator (S). Indeed in this case, the deterministic part of the stochastic differential equation would become negligible with respect to its stochastic part and the model would behave like in the model from eqn (2). Like in the cointegration model, we will add $S_{l,t}$ in front of the stochastic part of the differential equation in order to enforce the non-negativity constraint on the price of S_l. The cointelation model is summarized by eqn (5). Note that, like in the cointegration model described by eqn (4), $\theta \in]0,1]$.

Figure 1: The cointelation model can be seen as the generalization of both the correlation and the cointegration models.



$$\frac{dS_t}{S_t} = rdt + \sigma dW_t$$
$$dS_{l,t} = \theta(S_t - S_{l,t})dt + \sigma S_{l,t}(\rho dW_t + \sqrt{1-\rho^2}\, dW_t^{\perp}) \qquad (5)$$

Here we will call ρ the correlation of the cointelation.

3 Intuition behind the cointelation test

3.1 Studying the rolling correlation

In order to introduce the concept of cointelated pairs, Damghani *et al.* [2] pointed out that in cointegrated pairs the rolling 10-day correlation could have very significant swings. This observation would obviously be true in cointelated pairs. However, performing a test with this idea as a basis would become difficult in situations where the recall force [in eqn (5)] is quite strong.

3.2 Studying the number of times cointelated pairs cross paths

Another potential suggestion could be to study how many times the normalized time series cross paths. For example, if l represents the length of our study, we are in a situation of cointelated pairs, and $\theta = 1$, then the number of times the time series would cross would be $\frac{l}{2}$. This is because, in eqn (5), $dS_{l,t}$ is calculated in such a way that $S_{l,t}$ "catches up" with S_t every time a gap between $S_{l,t}$ and S_t has been created, which happens one time out of two alternately.

3.3 Studying measured correlation on different time intervals

Although the "rolling correlation," as described in Section 3.1, appeared promising, it does have its limitations – especially at the boundaries of definition of the cointelation model. The methodology we are proposing aims to be self-sufficient no matter whether the model is close to its limits or not. We will now introduce the intuition. Pairs that are cointelated will have, in the expectation sense, measured correlation that will be bigger as the time scale for the measurement of returns increases. Also, pairs that are cointelated will have a higher acceleration toward a measured correlation of 1 when the recall force θ increases. These claims should be immediately intuitive when one examines eqn (5). However, what might not be necessarily intuitive is that the bigger the time scale of measurement, τ, the bigger the probability that our returns fall "out of phase."[7] For example, if we examine Figure 7 – two cointelated pairs (which we will prove in Section 7) – we can see that if we happen to choose our first day of estimation for instance where spreads are bigger than their historical mean (e.g., June 2008) then, because of the recall force, we start our measured correlation process "out of phase" and we get a small correlation because the first few measurements go in the direction opposite to what the true long-term relationship would suggest. As we explain in Section 5, we would like to test whether the stochastic processes are cointelated. We cannot conveniently choose where to pick the first day of our measurement, so we will have to choose random first days of measurement, increase τ iteratively, and look whether on average this yields an increase in the measured ρ_τ.

4 Inferred correlation and term structure of correlation

For this section we have simulated a few cointelated pairs following eqn (5). We analyze our observation, and show that the concept of cointelation can be mapped onto the concept of inferred correlation at different time gaps.

4.1 Observations

We can see in Figure 2 that, depending on the chosen time scale for measurement, with a correlation of cointelation $\rho = -1$, we can actually achieve the whole span of possible measured correlations $]-1,1[$. We can also see that at the smallest time scale τ, the measured correlation happens to be equal to the correlation of cointelation ($\rho = \mathbb{E}[\hat{\rho}_1]$) as described by eqn (5). Furthermore, the variance of the measured correlation converges in the same figure.

4.2 Term structure of correlation

Intuitively, and through examination of the various figures (for example, Figure 2), we can see that the function which models the variation of measured correlation with respect to the time gaps between measurements should be an increasing function starting at $\tau = 1$, should converge toward 1 when τ goes toward ∞, and should go toward 1 faster when the recall force θ is bigger. One such function is described by eqn (6). In eqn (6) we can see that when τ goes toward 1, $e^{-\theta(\tau-1)}$ goes toward l and therefore $\mathbb{E}[\rho_t]$ goes toward ρ. On the contrary, when τ goes toward ∞, $e^{-\theta(\tau-1)}$ goes toward 0 and $\mathbb{E}[\rho_t]$ goes toward $\rho + (1-\rho) \times 1 = 1$.

$$\mathbb{E}[\rho_\tau] \approx \rho + (1-\rho)[1 - e^{-\theta(\tau-1)}],\ \tau \in \mathbb{Z}^*,\ \theta \in [0,1] \quad (6)$$

5 Cointelation test

In this section we define what makes two stochastic processes cointelated in three steps. First, we introduce Lemma 5.1 in order to formalize the intuition of "increasing measured correlation" described in Section 3.3, and then we formalize via Lemma 5.1 the intuition of the "number of crosses" idea given in Section 3.2.

5.1 Test formalization

Lemma 5.1 *The non-linear best fit of* $\mathbb{E}[\sup_{0<t\le\tau} \rho_t]$ *can be coherently modeled by eqn (7).*

$$\boxed{\mathbb{E}[\sup_{0<t\le\tau} \rho_t] \approx \rho + (1-\rho)[1 - e^{-\lambda\theta(\tau-1)}],\ \tau \in \mathbb{Z}^*, \theta \in [0,1], \lambda > 0} \quad (7)$$

Proof. We set $f(y_t,t) = y_t e^{\theta t}$

where: $\begin{cases} dx_t = x_t \sigma dW_t \\ dy_t = \theta(x_t - y_t)dt + \sigma y_t(\rho dW_t + \sqrt{1-\rho^2}\, dW_t^\perp) \end{cases}$

Differentiating, we get $df = \theta e^{\theta t} dt + e^{\theta t} dy_t$. Substituting for dy_t, we get $df = \theta e^{\theta t} x_t dt + e^{\theta t} \sigma y_t(\rho dW_t + \sqrt{1-\rho^2}\, dW_t^\perp)$. Now integrating, we get $y_t e^{\theta t} = \int_0^t \theta e^{\theta s} x_s ds + \int_0^t e^{\theta s} \sigma y_s(\rho dW_s + \sqrt{1-\rho^2}\, dW_s^\perp) - y_0$. We then get $y_t = x_t + (y_0 - x_0)e^{-\theta t} + e^{-\theta t} \int_0^t e^{\theta s} \sigma y_s ds(\rho dW_s + \sqrt{1-\rho^2}\, dW_s^\perp)$. Calculate $cov(y_t,y_t)$, $cov(x_t,y_t)$, and $cov(x_t,x_t)$ in order to plug into $\rho_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$. Substituting in, taking advantage of the fact that $< dW_t, dW_t^\perp >= 0$ and Ito's lemma, we get after integrating eqn (7).[8]

Lemma 5.2 *If l happens to be the length of the data, the expectation for the number of times the best chosen normalized returns of our cointelated pairs x and y cross paths,* $\mathbb{E}[\Gamma_{x,y}(\theta, l)]$, *should be equal to the expectation of the number of times correlated pairs cross paths,*[9] γl, *if* $\theta = 0$ *and should be equal to* $\frac{l}{2}$ *when* θ *converges to* 1. *The equation that best fits this claim empirically is eqn (8).*

$$\boxed{\mathbb{E}[\Gamma_{x,y}(\theta, l)] \approx l[\gamma(1-\theta) + \frac{1}{2}\sqrt{\theta}]} \quad (8)$$

Proof. We have the function $\mathbb{E}[\pi_{x,y}(l)] = \mathbb{E}[\Gamma_{x,y}(0,l)] = \gamma l$, $\mathbb{E}[\Gamma_{x,y}(1,l)] = \frac{l}{2}$. For small γ, $\mathbb{E}[\Gamma_{x,y}(\varepsilon, l)] \approx \gamma \frac{l}{2}\sqrt{\theta}$. The speed at which $\mathbb{E}[\Gamma_{x,y}(\theta, l)]$ increases should be proportional to the speed at which the variance increases for an $\mathbb{N}(0,t)$ -distributed random variable. As such, $f(x) = \sqrt{x}$ represents a good candidate.

Definition Two stochastic processes aimed at representing financial data will be cointelated if:

- Lemma 5.1 is verified.
- Lemma 5.2 is verified.
- The underlying assets have a reasonable physical connection that would suggest their spread should mean revert.
- In instances where the first two bullet points are not verified exactly, the correlation model cannot possibly be a substitute as correlation is a special case of cointelation (where $\theta = 0$).

Note that the third bullet point seems to be a rather vague statement relative to what one may expect in a usual mathematical definition. This rather unexpected approach aims to raise awareness of the fact that this remains a model for which the application should be properly understood, rather than being blindly used – which is so often the case in the financial industry.

5.2 Test verification

Figure 3 represents a comparative study of cointelated and correlated pairs through simulation of a realistic financial θ. As one can see, $\mathbb{E}[\sup \rho_\tau]$ fits well $\max \hat{\rho}_\tau$. In the case where our recall force θ is very weak, the $\mathbb{E}[\sup \rho_\tau]$ discussed in Lemma 5.1 will represent a lower bound rather than an expectation. Figure 4 illustrates this particular property. In the case θ is strong, $\mathbb{E}[\sup \rho_\tau]$ discussed in Lemma 5.1 will represent a slight upper bound. In these cases the pairs are so obviously cointelated that the test is not really needed. Figure 5 illustrates this idea.

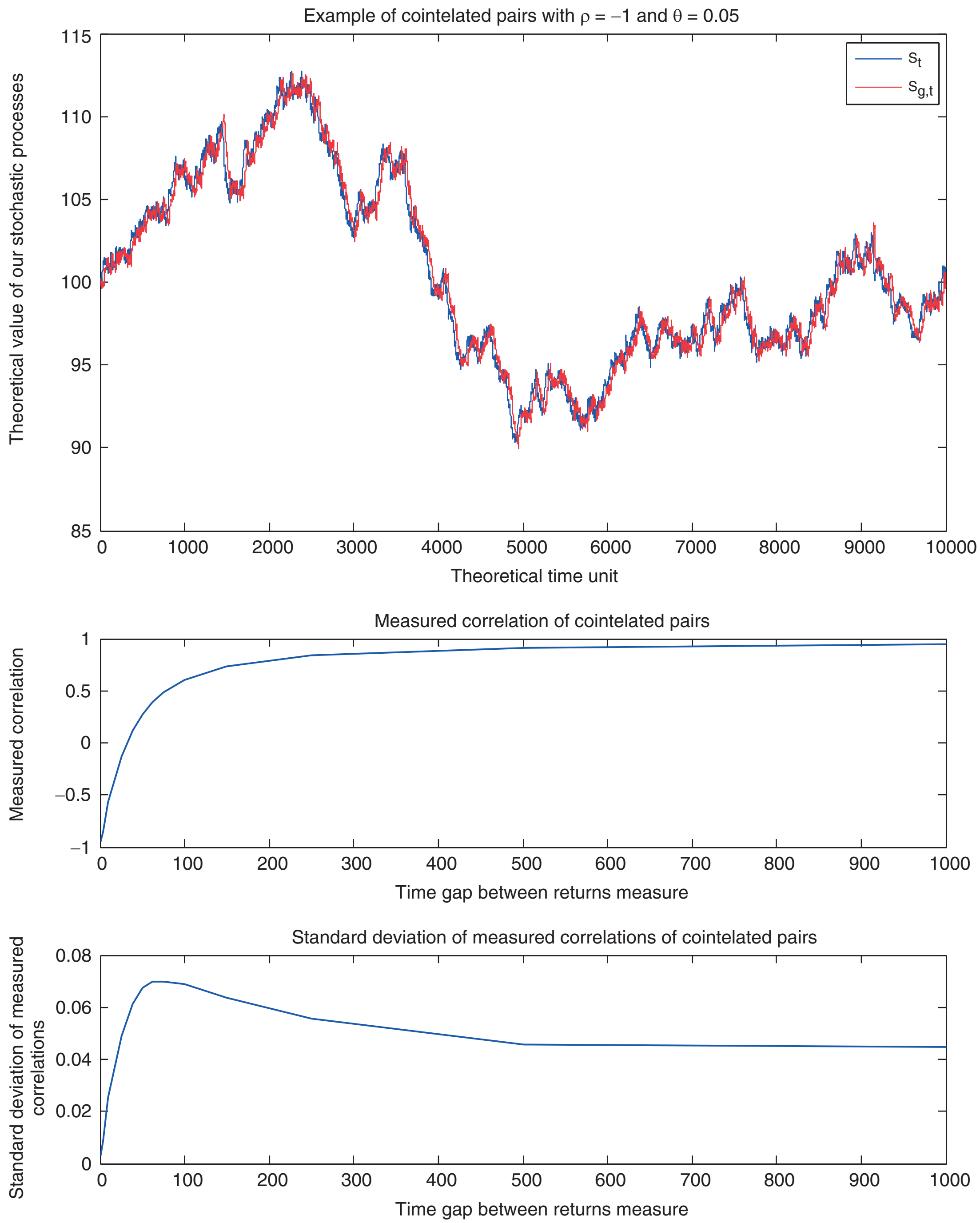
6 Note on parameter estimation

6.1 Estimating θ by rearranging the SDE

Note that $\hat{\rho}_1 = \mathbb{E}[\rho_1]$. By rearranging eqn (5), we get

$$\hat{\theta} = \mathbb{E}[\frac{dS_{l,t}}{(S_t - S_{l,t})dt + \sigma S_{l,t}(\hat{\rho}_1 dW_t + \sqrt{1-\hat{\rho}_1^2}\, dW_t^\perp)}]$$, and therefore

Figure 2: Example of cointelated pairs spanning all the measured correlation spectrum conditional on the assigned time scale.

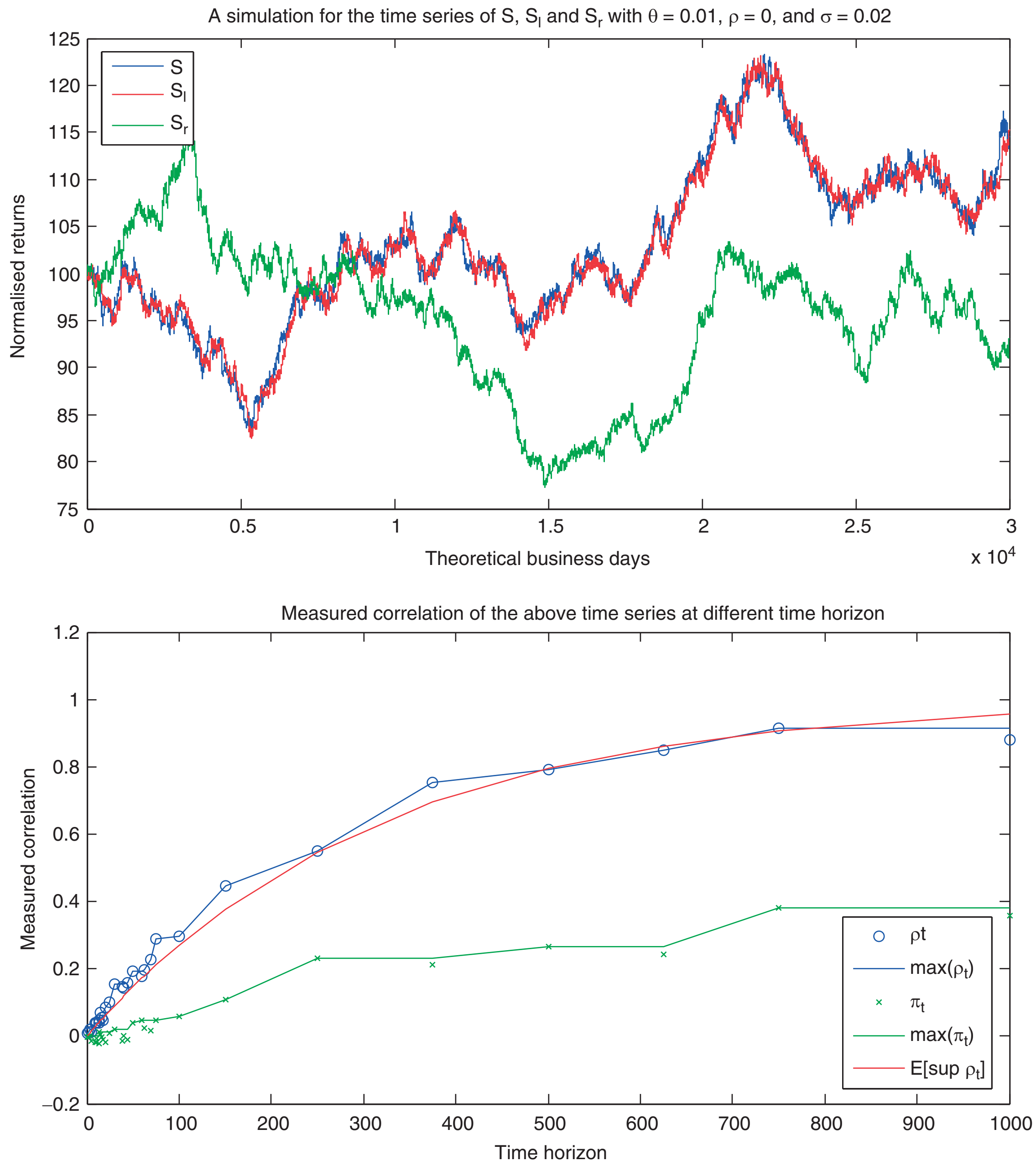


$\hat{\sigma} = \mathbb{E}[\frac{dS_{l,t} - \hat{\theta}(S_t - S_{l,t})dt}{(\hat{\rho}_1 dW_t + \sqrt{1-\hat{\rho}_1^2}\, dW_t^{\perp})}]$. Similarly to the variance reduction methodology described in [2], we will define $B_+ = |\frac{\max(S_t - S_{l,t}, t \in [0,T])}{2}|$ and $B_- = |\frac{\inf(S_t - S_{l,t}, t \in [0,T])}{2}|$. We note that the estimation of θ is noised when $Z_\sigma = B_+ > |S_t - S_{l,t}| > B_-$ where σ, in contrast, has quality samples. The reverse is true when $Z_\theta = |S_t - S_{l,t}| > B_+ \bigcup |S_t - S_{l,t}| < B_-$. We will therefore sample θ in Z_θ and σ in Z_σ. Figure 6 gives a representation of these sampling zones.

6.2 Estimating θ via the inferred correlation formula

Damghani and co-workers [2, 5] showed a way to reduce the variance for the θ parameter in the cointelation model (or more generally in the OU process infamous for being slow to converge[10] in the industry). From eqn (6), it is tempting to rearrange

Figure 3: Comparative study of cointelated and correlated pairs through simulation of a realistic financial.



the equation so that an additional estimator for θ is isolated. The solution from the rearrangement is given by

$$\theta \approx \frac{1}{n}\sum_{i=1}^{n}[-\frac{1}{\lambda n}ln(\frac{\max(\hat{\rho}_n)-1}{\rho-1})] \qquad (9)$$

where n happens to be the biggest time gap for which $\hat{\rho}_n$ happens to have enough samples to make it statistically significant (around 30).

Proof. Rearranging eqn (6), we get $\hat{\rho}_\tau - \rho = (1-\rho)(1-e^{-\lambda\theta\tau})$. Taking the log on each side, we get $ln(\frac{\hat{\rho}_\tau-\rho}{\rho-1}+1) = -\lambda\theta\tau$. Rearranging further and taking the expectation, we get eqn (9).

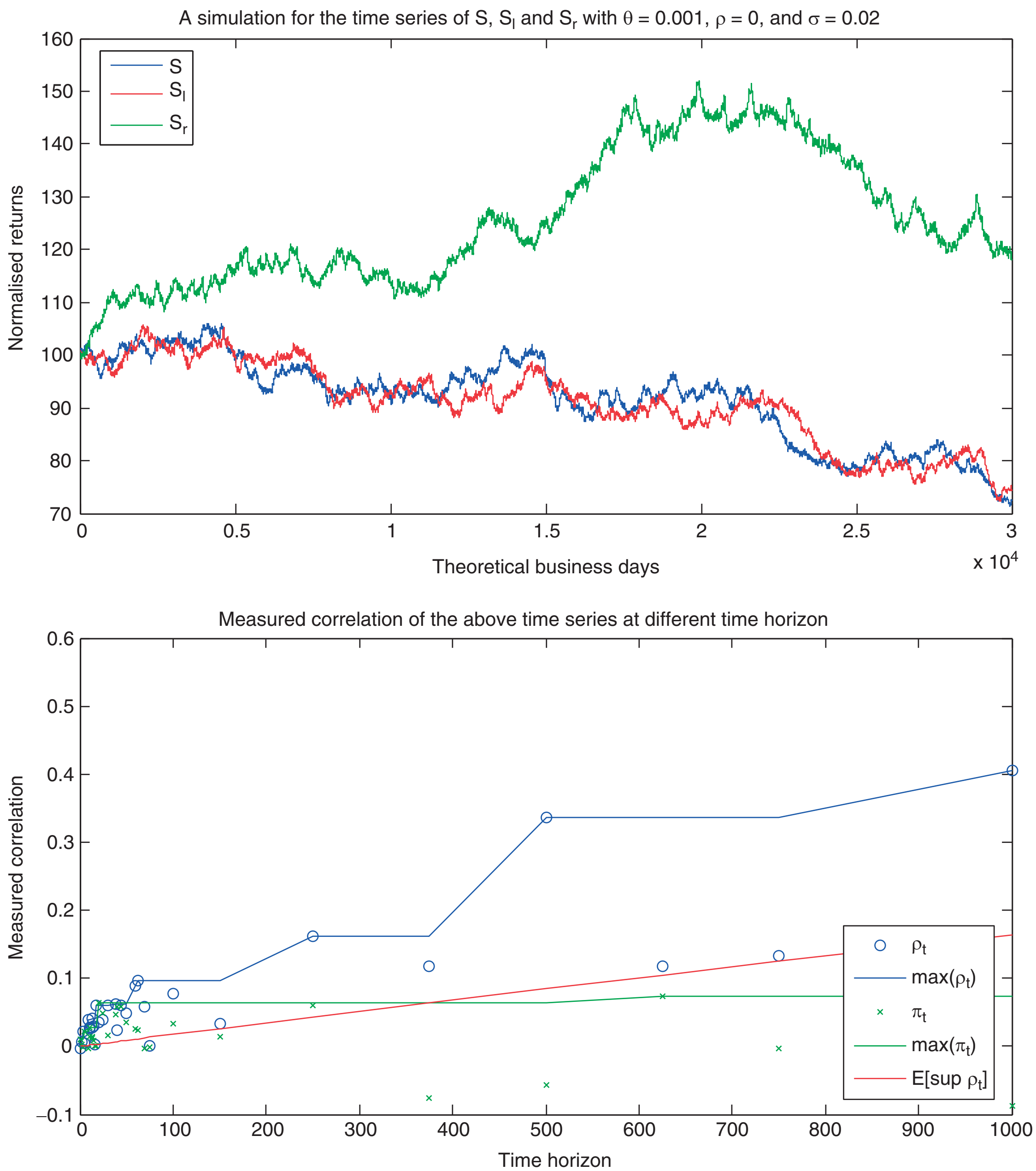
Although tempting, this methodology usually yields a poor estimator for θ. This is because, although eqn (9) pretty accurately predicts the general shape of max $\hat{\rho}_\tau$, it however remains an estimation whose error in [0,1] is enhanced by the log function in eqn (9).

6.3 Estimating θ via the number of crosses formula

As we can rearrange eqn (6) to isolate and get an estimation of θ in eqn (9), we can do the same with eqn (8). Indeed, another estimation of θ can be given by

$$\theta \approx \lambda(\frac{\hat{\Gamma}_{x,y}(\theta,l)}{l}-\gamma)^2 \qquad (10)$$

Figure 4: Comparative study of cointelated and correlated pairs through simulation of a very weak.



Proof. By rearranging $\hat{\Gamma}_{x,y}(\theta, l) = l(\gamma + \frac{1}{2}\sqrt{\theta})$ we get eqn (10).

6.4 Variance reduction technique for θ

Combining the variance reduction methodology described in [2, 5], we can find an estimator for $\theta(\hat{\theta}_d)$. Combining the results from eqns (9) and (10), we can find a better estimator for θ as given by

$$\theta \approx w_1\hat{\theta}_d + w_2\left(\frac{1}{n}\sum_{i=1}^{n}\left[-\frac{1}{\lambda n}ln\left(\frac{\hat{\rho}_n - 1}{\rho - 1}\right)\right]\right) + (1 - w_1 - w_2)\lambda\left(\frac{\hat{\Gamma}_{x,y}(\theta, l)}{l} - \gamma\right)^2 \qquad (11)$$

where w_1, w_2, and $(1 - w_1 - w_2)$ are in [0,1] and represent the optimal weights, whose value will be addressed in a future paper.

7 Application to the example of oil and BP

7.1 Physical reason

Because BP is a company focusing on the production of oil, it is normal that its long-term performance is strongly linked to that of oil. This means that the relationship between oil and BP is a good candidate for the cointelation model.

7.2 Correlation on different time scales

If one examines the correlation between oil and BP (Figure 7), one realizes that the measured correlation between oil and BP is a function of the gaps in the measurement of the returns. For example, if we measure the correlation of the returns

Figure 5: Comparative study of cointelated and correlated pairs through simulation of a very strong.

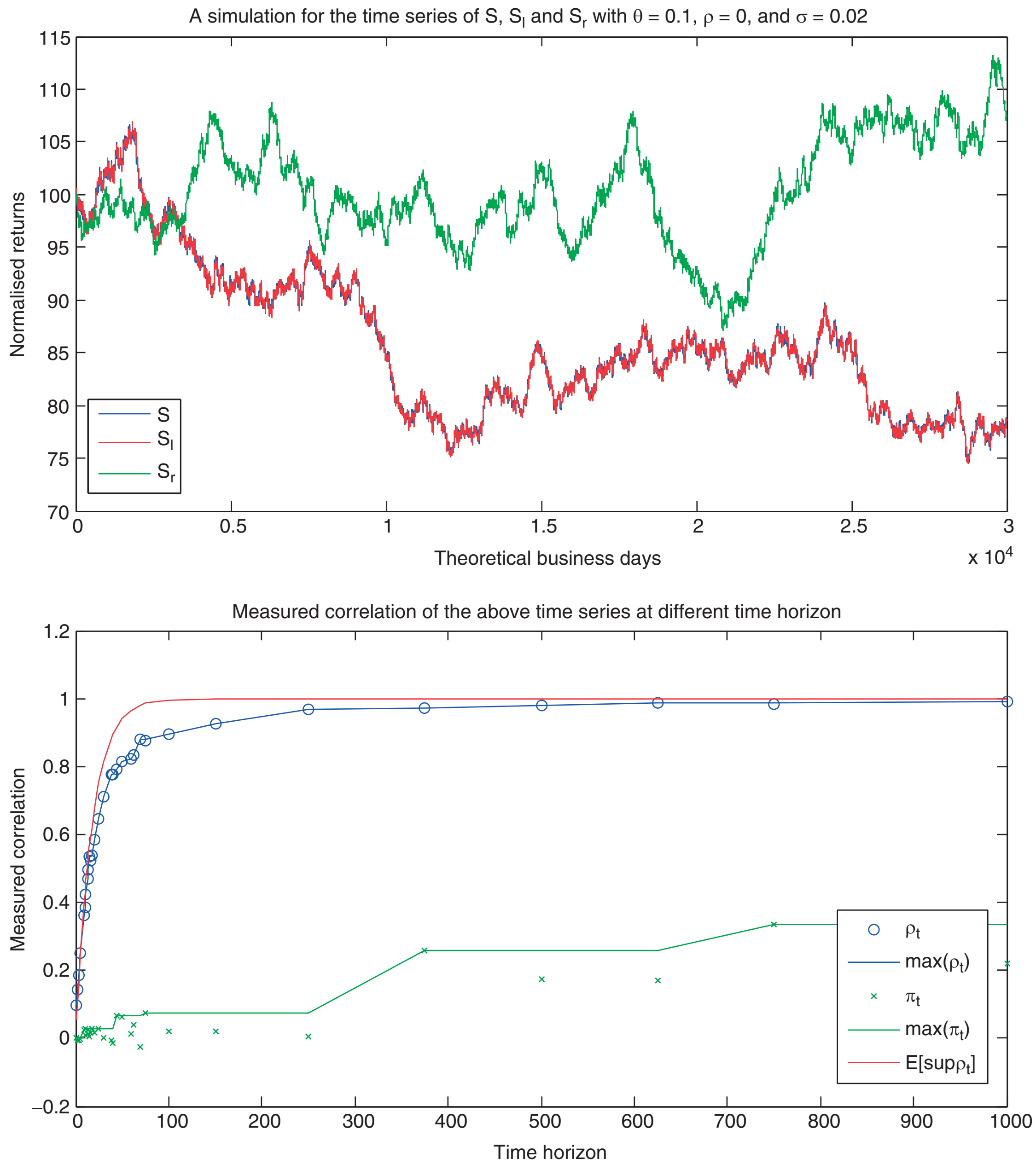
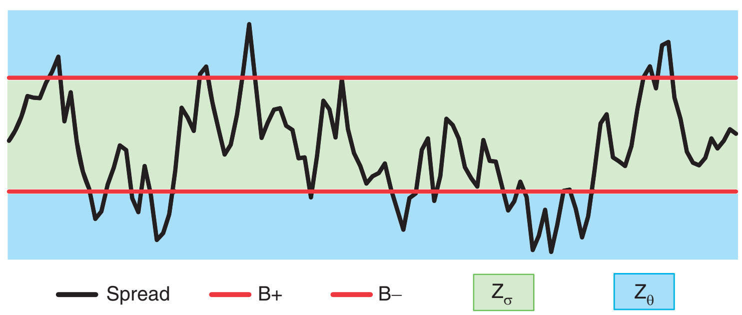
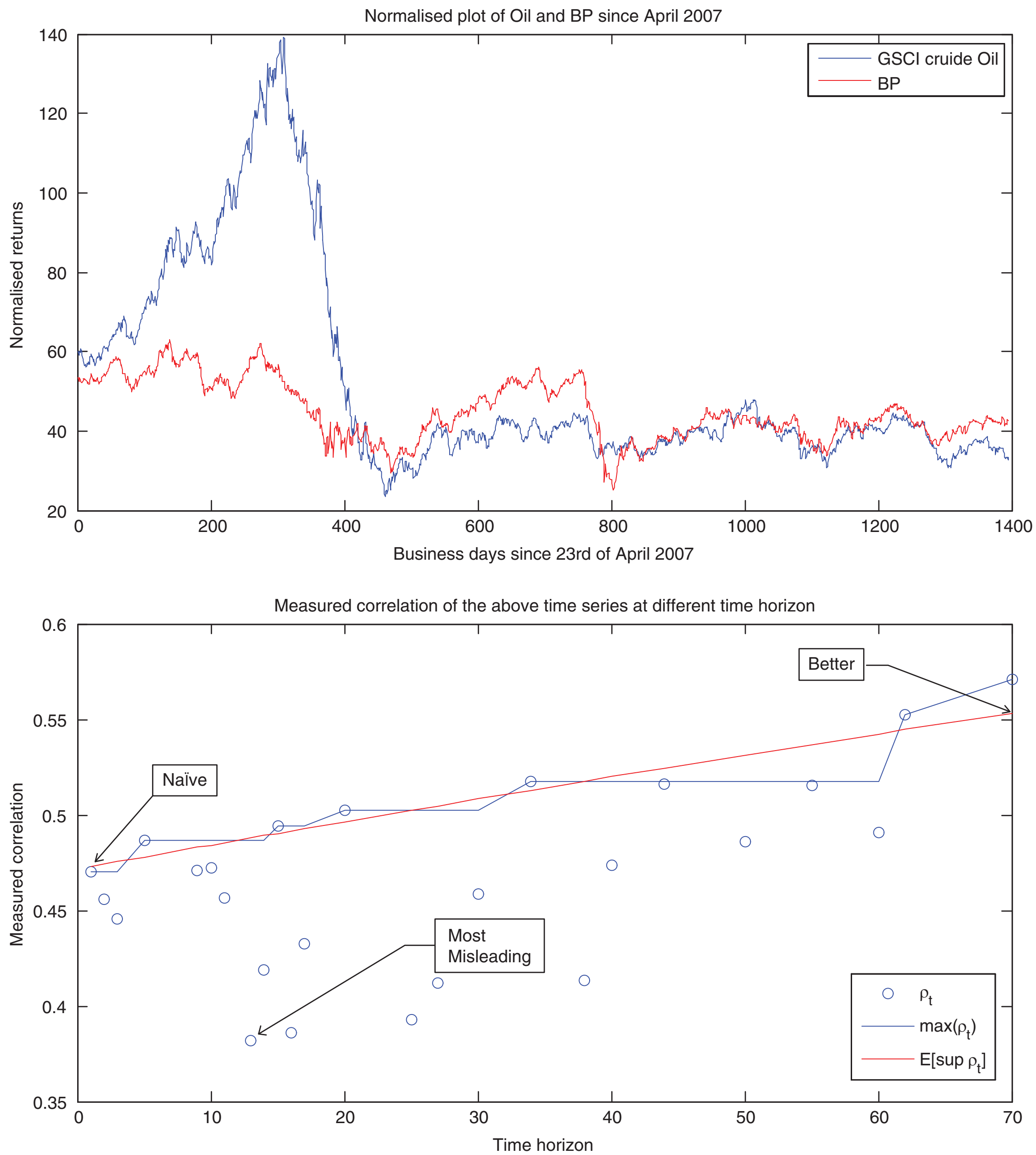


Figure 6: Visual representation of the sampling zones for θ and σ.



between oil and BP every 12 days, we get around 0.37 whereas if we do it every 70 days, we get a correlation of 0.56. How do we explain this discrepancy? The difference is due to the fact that in the long run, oil and BP have the same drift but in the short run, we get a weaker perceived relationship due to the numerous noises that can only impact one of these assets and not the other (for example, the very temporary impact of the Gulf of Mexico accident on the spread between oil and BP in Figure 7). For the physical reasons just described, the relationship between oil and BP is a good candidate for the cointelation model. In fact, with the current financial mathematics literature this duality between differences in the short-term relationship and the long-term relationship can only be explained by the cointelation model. Using the variance reduction technique [2, 5], one can estimate θ (in the last 5 years,

Figure 7: Time series of oil and BP in the last 5 years and statistics about their measured and inferred correlation estimates.



for daily returns $\hat{\theta} \approx 1\%$) and then plug θ into eqn (7) and we get a pretty accurate representation for $\mathbb{E}[\max \hat{\rho}_\tau]$ (Figure 7). Figure 8 shows that a similar study can be done with Target and Walmart.

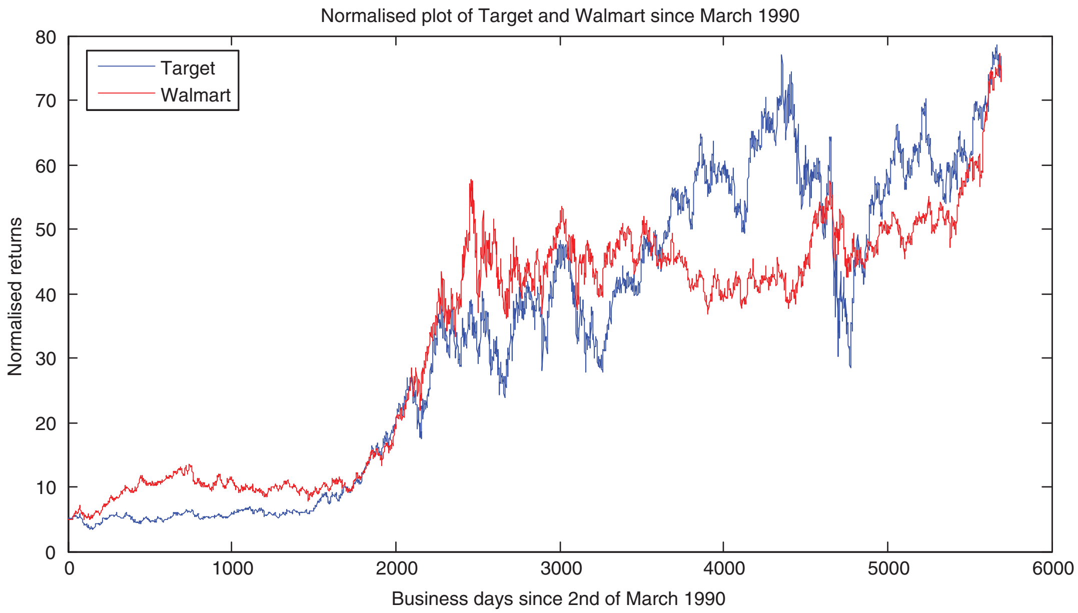
7.3 Number of crosses

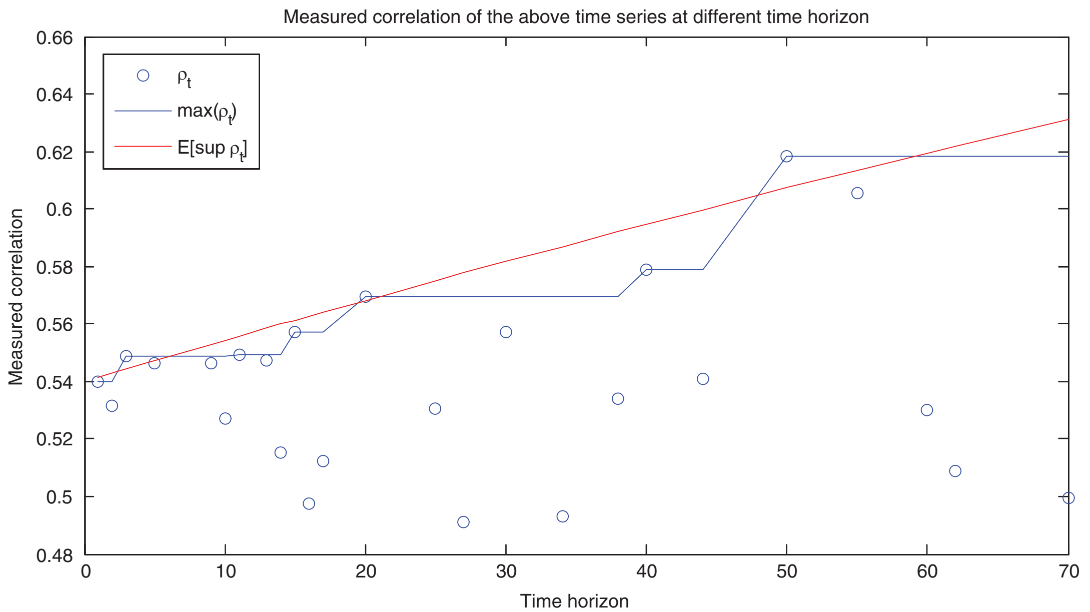
If one wishes to argue that the difference between 0.56 and 0.37 is not that significant, then one can also examine the number of crosses $\hat{\Gamma}_{x,y}(\hat{\theta}, l)$, which corroborates the result described in Section 7.2. Indeed, plugging $\hat{\theta} = 1\%$ in eqn (8), we get $\hat{\Gamma}_{x,y}(\hat{\theta}, l) \approx \frac{3.2l}{100}$ (where $\mathbb{E}[\Gamma_{x,y}(\hat{\theta}, l)] \approx \frac{5.5l}{100}$ and $\mathbb{E}[\Gamma_{x,y}(0, l)] = \mathbb{E}[\Pi_{x,y}(l)] \approx \frac{l}{100}$).

7.4 Prediction and biological explanation

Because there is not enough data to measure the longer-term correlation, one can use eqn (9) and find the yearly correlation estimate which gives $99\% = 0.47 + (1 - 0.47)(1 - e^{-1.75(1\%)(254-1)})$, a very strong long-term relationship. Note that this value seems very high. The reason for this perception is twofold. First, this gives an estimate of $\mathbb{E}[\max \hat{\rho}_\tau]$ which should be equivalent to $\mathbb{E}[\hat{\rho}_\tau]$ provided the right phase is chosen. The probability that the right phase is chosen decreases as τ increases. However, the longer the time scale, the smaller the penalty for being "out of phase." The second reason is more qualitative and inspired by the UTOPE concept introduced by Gosset [4]. Indeed, because the benefit of seeing a true pattern

Figure 8: Time series of Target and Walmart in the last 22 years and statistics about their measured and inferred correlation estimates.





far outweighs the cost of being mistaken with respect to false patterns, our minds are automatically set to see patterns sometimes when there are none. The same way it was difficult for our ancestors to grasp that evolution occurs since we cannot physically "see" (with the naked eye) evolution occurring, we might not "see" why correlation should increase that much. This is because the time scale over which we have evolved is the product of environmental pressures that did not select for these "long-term" qualities which would make us grasp the differences of time scale. We just do not live long enough. You cannot "see" evolution with the naked eye, you acknowledge it through evidence. Similarly, seeing the pattern that the correlation is low on a time scale we are comfortable with makes extrapolation difficult. Careers in finance rarely last more than 30 years. Had careers lasted 300 years, these phenomena would be much more intuitive. It might not be very intuitive at first to see why the long-term correlation we have never really "seen" is high,[11] but the mathematical evidence suggests the opposite.

8 Application to the financial industry

There are plenty of places where inferred correlation or cointelation could be used. We will now examine a few of them.

8.1 Extrapolating daily correlation from monthly marked correlation

Traders understand that measured correlation does not accurately represent the relationship between assets in the long term as it fails to capture the long-term drift

Figure 9: Relationship between marked correlation and the ratio of the relevant underliers.

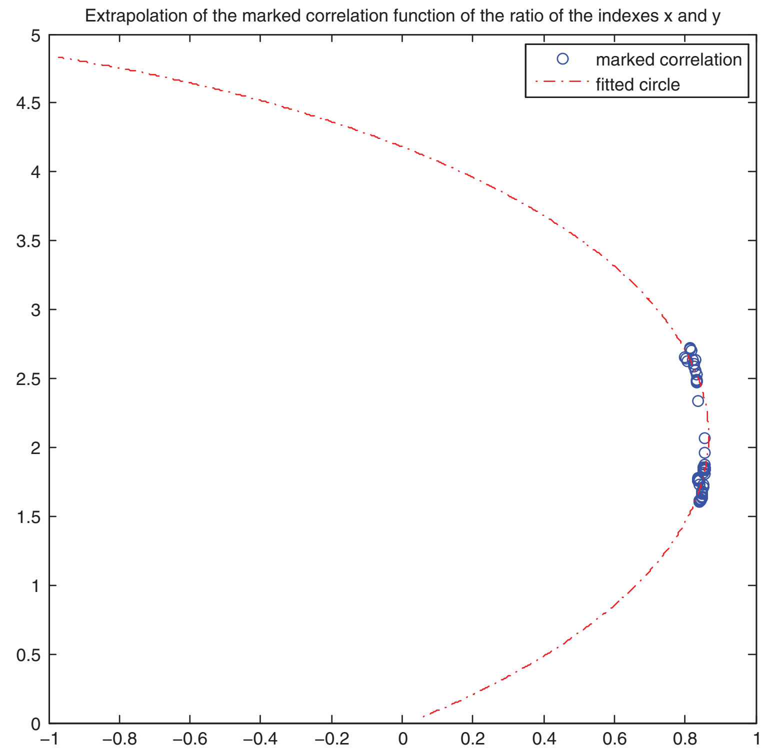
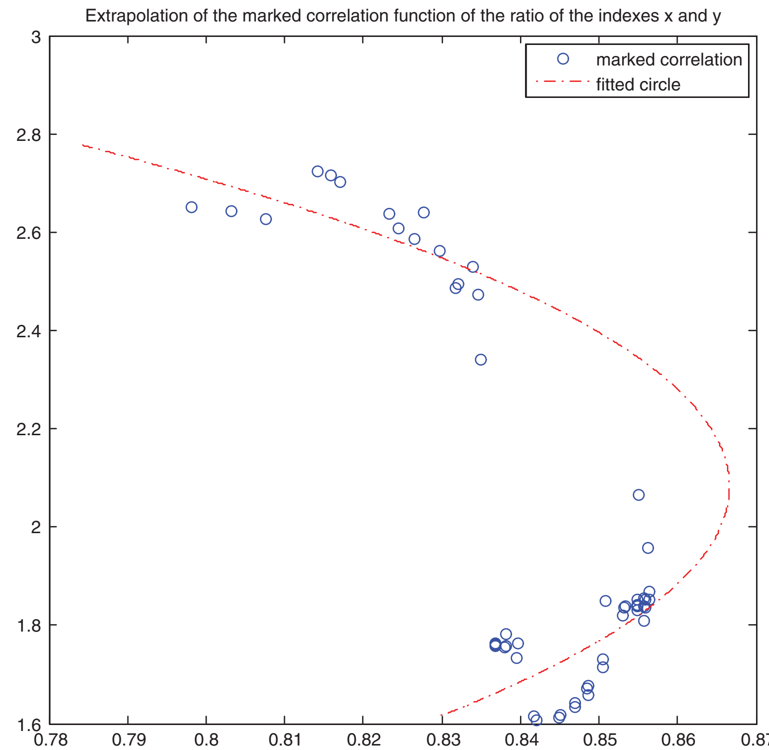


Figure 10: Relationship between marked correlation and the ratio of the relevant underliers zoomed in.



of the underliers. As such, traders understand that in situations where the perceived equilibrium between assets is away from its historical mean, the expectation of future realized correlation should be smaller than its historical mean correlation. A way to represent this concept mathematically is via the equation of a semicircle. Indeed, if we call Ω the set of points (x, y) available for marketed correlation, then the semicircle that will best fit these market correlation can be retrieved using

$$\hat{x}_c, \hat{y}_c, \hat{r} = \arg\min_{x_c, y_c, r} \sum_{i=1}^{N} [(x + x_c)^2 + (y + y_c)^2 - r^2], (x, y) \in \Omega \qquad (12)$$

Note that Figure 9 and its zoomed version, Figure 10, support the idea that traders believe underliers have the same drift in the long run. Indeed, these figures suggest that when the ratio of the underlier x over the underlier y is away from some sort of natural mean, then in anticipation that the underliers should mean revert the traders mark an implied correlation smaller than if the ratio was at its mean.

8.2 Measuring risk in the long and in the short term, and extrapolating correlation in time scales otherwise unmeasurable

We have shown that the cointelation model was able to model both the short- and long-term risk associated with two underliers. This has obvious direct consequences in measuring VaR at different time lapses with the same model. If a risk manager would like to assess the long-term relationship between two assets but does not have enough data to do so, he/she can use the concept of inferred correlation. For example, if the risk manager would like to assess the yearly correlation between a commodity and its equity mirror but this specific equity does not have enough data, the risk manager can use daily returns in order to estimate θ from the cointelation model and then extrapolate the long-term correlation thanks to eqn (6) and work with a parameter that is far easier to handle and more popular than the cointelation model, which would represent a more complete and accurate relationship between the two underliers. Also, any options involving a correlation should try to link the "marked" correlation with respect to an inferred correlation rather than widening the bid–ask spread of the priced option for which the underliers appear to have the properties of cointelated pairs.

8.3 Stressed VaR and capital requirement

The capital requirement as described by Basel III [6] mentions: "*Going forward, banks must determine their capital requirement for counterparty credit risk using stressed inputs. This will address concerns about capital charges becoming too low during periods of compressed market volatility and help address* ***procyclicality***." Procyclicality is a fancy term that essentially means correlations tend to 1 with the economy, especially in periods of large movement. Although the concern of the Basel committee happens to be a fair observation of the markets going forward, the translation of this concern into policy needs to be adjusted. Indeed, the capital requirement formula [eqn (13)] as currently understood is that

$$CR = k \times (VaR + SVaR) \qquad (13)$$

where CR represents the capital requirement of the financial institution, k is a number reflecting the perceived solidity of the bank, *VaR* is the value at risk, and *SVaR* is the stressed value at risk. However, a simple observation and reflection shows that the capital requirement as defined by this formula could yield potentially an output superior to the value of the total assets being stress tested to begin with.

This capital requirement is therefore overly punitive as it prevents many desks from pursuing their businesses while at the same time not necessarily addressing the procyclicality of the market, as it aims to achieve. At first glance it would appear more logical to set the capital requirements as per

$$CR = ESF_{T-\tau} \quad (14)$$

where $ESF_{T-\tau}$ would represent the expected shortfall under the inferred correlation hypothesis in the time lapse going from today (T) until the beginning of the last crisis (τ).

8.4 Portfolio construction

Correlation is a central parameter of Markowitz's modern portfolio theory [7] and the basis for the business model of many hedge funds and asset managers. However, it is not really clear whether the limitations of the correlation model are properly understood on the theoretical side, and certainly in the application of this theoretical side. Also, if one wishes to pair trade two underliers for which the historical mean has been modified by either technology or physical reasons, one may wish to study the evolution of measured correlation based on different time lapses in order to determine what the micro business cycle associated with this pair is. Obtaining this information may lead to better-thought-out mean-reverting or dispersing trading strategies.

8.5 Marketing material and proper market conduct

We have seen that, in the example of oil and BP, the measured correlation was very dependent on the time lapse of measurements τ. It could therefore be tempting for a sales person to try to capitalize on the lowest value of this measured correlation. For example, a commodities sale could advertise the lowest measured correlation available to try to attract long-term investors who usually invest in the equities market and who seek diversification in the long term. These investors do not necessarily understand that the measured correlation is an increasing function of τ in cointelated pairs. These investors, who then adjust their portfolio based on the theory laid down by Markowitz [7], would get an overall risk in their portfolio which would not go toward their diversification strategy. These practices are in violation of the **Client Best Interest Rule** and the rule on **Misleading Statement and Actions** laid out by the FCA. For example, in the bottom graph of Figure 7, which represents the measured correlation between BP and oil at different timescales, a negligent, naïve salesperson may want to return the measured correlation using daily returns (here around 0.46) in order to justify to the long-term investor that the relationship between oil and BP is weak. Could a salesperson decide to do worse than that? He could in fact write a for loop in which he will calculate the correlation at different timescales and decide to return the correlation which measure is the smallest possible. In Figure 7, that falls on a correlation of around 0.36 using 12 days return, which does not fall onto a natural symbolic timescale, i.e., daily, weekly, monthly, or yearly return. For correlations that are potentially cointelated, or if there is dispute on which measure correlation to choose, a salesperson should always take the conservative approach for marketing material. In this example, if he defines a long-term investor as an investor who rebalances his portfolio every 70 days, he should use the formula given by the inferred correlation formula (here the red line points to something around 0.55).

9 Conclusion

We have summarized the three models introduced by Damghani et al. [2] and given the intuition behind the cointelation test, which uses the increasing value of measured correlation at different time scales as well as the concept of number of crosses. We have also explained the connection between the cointelation stochastic process and the concept of inferred correlation, whose objective is to map the properties of the cointelated pairs onto an inferred correlation estimate that people are more familiar with. Finally, we have illustrated our findings through the example of oil vs. BP and Walmart vs. Target, as well as stated a few applications of this work in the financial industry.

Babak Mahdavi Damghani is an FSA-approved person who has worked at Socgen, Cantab Capital Partners, Citigroup, and Credit Suisse in, respectively, Equities Derivatives High Frequency Statistical Arbitrage, FX Quantitative Strategies, Exotics Derivatives Commodities Trading, Hybrid Structuring, and Cross Asset Quantitative Risk Methodology. He has postgraduate degrees from Oxford, École Polytechnique, and Cambridge in various mathematical and computational sciences.

ENDNOTES

1. The choice of the title refers to "The misleading value of measured correlation," which this paper is a sequel to.
2. D. Sornette. Power laws and scaling in finance: Practical applications for risk control and management. Available online: http://www.er.ethz.ch/presentations/PowerScaling_risk.pdf
3. B. Mandelbrot. Fractals and scaling in finance: Discontinuity, concentration, risk. 1997.
4. B.M. Damghani, D. Welch, C. O'Malley, and S. Knights. The misleading value of measured correlation. 2012.
5. Please note here that the term cointegration is perhaps slightly differently formulated from the usual cointegration models presented in the econometrics literature. The term cointegration here represents a technical jargon introduced in "The misleading value of measured correlation," whose aim is to specify the concept of mean reversion.
6. Presented in a slightly different manner in "The misleading value of measured correlation."
7. Also, the bigger the time scale the less data we will have and the more biased our data will be, which does not help the estimation of correlation.
8. $\lambda \approx 1.75$ for "regular" financial data. In reality, λ is itself a function of the other parameters. This concept will be developed more rigorously in a subsequent paper.
9. With reasonable financial data, $\gamma \approx 0.01$. Like λ, γ is itself a function of the other parameters. This concept will be developed more rigorously in a subsequent paper.
10. Please note that this is different from suggesting that the relationship is weak.
11. Or rather, we see it during the crisis in which the correlation of everything goes to 1, but prefer to interpret it differently.

REFERENCES

[1] I. Gosset. Utope-ia. *Wilmott*, Issue 60 (2012), pp. 28–37.
[2] B.M. Damghani, D. Welch, C. O'Malley, and S. Knights. The misleading value of measured correlation. *Wilmott*, Issue 62 (2012), pp. 64–73.
[3] P. Wilmott. *Frequently Asked Questions in Quantitative Finance*, 2nd edn. John Wiley & Sons, New York, 2009; pp. 336–337.
[4] G.E.Uhlenbeck and L.S.Ornstein. On the theory of Brownian motion. *Physical Review*, Vol. 36 (1930), pp. 823–841.
[5] B.M. Damghani and A. Kos. De-arbitraging with a weak smile. *Wilmott*, Issue 64 (2013), pp. 40–49.
[6] Basel Committee on Banking Supervision. Basel III: A global regulatory framework for more resilient banks and banking systems (2003), p. 11.
[7] H. Markowitz. Portfolio selection. *The Journal of Finance*, Vol. 7, No. 1 (1952), pp. 77–91.