On Epps Effect and Rebalancing of Hedged Portfolio in Multiple Frequencies

Mustafa U. Torun and Ali N. Akansu
Department of Electrical and Computer Engineering
New Jersey Institute of Technology
University Heights, Newark, NJ 07102 USA
mustafa.torun@njit.edu
akansu@njit.edu

Abstract—Correlations of financial asset returns play a central role in designing investment portfolios by using Markowitz’s modern portfolio theory (MPT). Correlations are calculated from asset prices that happen at various trading time intervals. Therefore, trading frequency dictates correlation values. This phenomenon is called the Epps effect in finance. We present variations of correlations as a function of trading frequency to quantify Epps effect. The results reiterate that portfolio variations of correlations as a function of trading frequency, phenomenon is called the Epps effect [6]. In this paper, we calculate correlations of asset returns decrease as the sampling (trading) period of prices decreases, and this phenomenon is called Epps effect [6]. In this paper, we calculate correlations of several assets and bring additional insights to this important phenomenon for improved hedging in multiple trading frequencies.

I. INTRODUCTION

Correlations of asset returns in an investment portfolio is an important aspect of MPT [1]. It also serves for various hedging and pairs trading strategies. A good estimation of correlation is crucial for good performance in all trading and risk management systems [2]. However, a good correlation estimation, especially in intra-day and high-frequency trading where sampling periods are typically below a minute, is of a major challenge [3–5]. It is known in finance that the correlations among financial asset returns decrease as the sampling (trading) period of prices decreases, and this phenomenon is called Epps effect [6]. In this paper, we calculate correlations of several assets and bring additional insights to this important phenomenon for improved hedging in multiple trading frequencies.

In Section II, we define pair-wise correlation of asset returns, and highlight its significance in finance. In Section III, we treat correlation as a function of trading interval with the help of Black-Scholes pricing model [7]. Then, we revisit the Epps effect and quantify the asynchronous trading as its usual suspect. We present closing remarks in Section IV.

II. CORRELATIONS OF ASSET RETURNS

Return of a financial asset over a time period of \( T \) is defined as

\[
R(n) = \frac{P(n)}{P(n-1)} - 1,
\]

where \( P(n) \) is the last traded price at discrete time \( n \) that is obtained by sampling the continuous-time last traded price of the asset at \( t = nT \) as given

\[
P(n) = P(t) |_{t=nT},
\]

It is customary to refer the asset returns with respect to their sampling periods, e.g. 30-min returns, 1-hour returns, end of day (EOD) returns. The cross-correlation coefficient of the returns of two assets, \( R_1(n) \) and \( R_2(n) \), is defined as

\[
\rho = \frac{E[R_1(n)R_2(n)] - \mu_1\mu_2}{\sigma_1\sigma_2}, \tag{3}
\]

where \( E[\cdot] \) is the expected value operator, \( \mu = E[R] \) is the mean, and \( \sigma^2 = E[R^2] - \mu^2 \) is the variance of a return process. Note that \( \text{expected return} \mu \), and \( \text{volatility} \sigma \) are the two widely used metrics in finance. Correlation between asset returns plays an important role especially in portfolio selection, hedging, and pairs trading. In Markowitz’s MPT [1], the optimum investment amount in each asset of an \( N \)-asset portfolio is obtained by solving the following minimization problem [2]

\[
q^* = \underset{q}{\text{argmin}} \frac{1}{2} q^T Cq + \lambda_1 (\mu_T - q^T \mu) + \lambda_2 (1 - q^T 1), \tag{4}
\]

where \( q = [q_1 \ q_2 \cdots \ q_N]^T \) is \( N \times 1 \) investment vector, \( q_i \) is the investment proportion of the total capital into the \( i \)-th asset, \( T \) is the matrix transpose operator, \( C = \begin{bmatrix} C_{ij} \end{bmatrix} = E[R_i(n)R_j(n)] - \mu_i\mu_j \) is the covariance matrix, \( \lambda_1 \) and \( \lambda_2 \) are two Lagrangian multipliers corresponding to the constraints \( q^T \mu = \mu_T \) and \( q^T 1 = 1 \), respectively, \( \mu_T \) is the target expected portfolio return, \( \mu = [\mu_1 \ \mu_2 \cdots \ \mu_N]^T \) is the \( N \times 1 \) expected return vector, \( \mu_i = E[R_i(n)] \) is the expected return for the \( i \)-th asset, and \( 1 \) is an \( N \times 1 \) vector with its elements all equal to one. Note that the covariance matrix \( C \), cross-correlations between returns of the assets in the portfolio, has a direct impact on the portfolio selection problem stated in (4).

Another application that the cross-correlation of asset returns comes into play is hedging. In this type of trading, returns of an asset in time are regressed on the returns of another asset (usually an exchange traded fund (ETF) that tracks the index of the corresponding industry) such that

\[
R_1(n) = \alpha + \beta R_2(n) + \xi(n), \tag{5}
\]

where \( \alpha, \beta, \) and \( \xi \) are commonly referred to as drift, systematic component, and idiosyncratic component, respectively. According to the model given in (5), return of the first asset is...
composed of a constant drift, $\beta$ times the return of the second asset, and white noise. The idea in hedging is to invest $1$ in the first stock and invest $-\beta$ in the second stock (hedge) such that the return on investment becomes

$$R_t(n) = R_1(n) - \beta R_2(n) = \alpha + \xi(n),$$

and the expected return on investment is calculated as $E[R_t(n)] = E[\alpha] + E[\xi(n)]$. This strategy is (i) profitable if $E[\alpha] > 0$ with $E[\xi(n)] = 0$ (ii) expected to profit if $E[\alpha] = 0$ and

$$X(n) = \sum_n \xi(n) \ll E[X(n)],$$

since $X(n)$ is expected to return to its mean, and expected to provide more positive samples than negative samples of $\xi(n)$. The latter is called pairs trading. It isolates the return on investment from the market (industry) and bets against the excess returns of a specific asset. By assuming $E[R_2(n)\xi(n)] = 0$ and substituting (5) into (3) we obtain

$$\beta = \sigma_1 / \sigma_2.$$  

(7)

It is seen from (6) and (7) that the performance of a pairs trading strategy is related to the cross-correlation coefficient of two assets, $\rho$. Therefore, better correlation estimation of asset returns is an important factor of good performance.

III. EPPS EFFECT

According to Black-Scholes price model [7] log-price, $S(n) = \ln P(n)$, of an asset follows a geometric Brownian motion path as expressed

$$S(n) = S(n-1) + G(n),$$

where $G(n) \sim N(\mu, \sigma^2)$ is the log-return. $G(n)$ is a stationary noise process and white, i.e.

$$E[G(n-k)G(n-l)] = \mu^2 = \sigma^2 \delta_{k-l},$$

where $\delta_{k-l}$ is the Kronecker delta function as given

$$\delta_{k-l} = \begin{cases} 1 & k = l \\ 0 & k \neq l \end{cases}.$$

Moreover, for small values of $R(n)$, return process of (1) is an approximation to the log-return due to the Taylor series expansion of the logarithm, i.e.

$$G(n) = \ln \left[ P(n)/P(n-1) \right] \approx [P(n)/P(n-1)] - 1 = R(n).$$

(10)

(11)

It is seen from (8) that we can write the log-price at discrete time $n$ as a sum of initial log-price and all log-returns up to $n$ as follows

$$S(n) = S(0) + \sum_{i=1}^{n} G(i).$$

Now, consider two discrete-time log-prices of the same asset, $S_{T_1}(n)$ and $S_{T_2}(n)$ sampled with sampling periods $T_1$ and $T_2$ according to (2), respectively. Let $T_2 = kT_1$. From (8) and (11) we write

$$G_{T_2}(n) = S_{T_2}(n) - S_{T_2}(n-1)$$

$$= S_{T_2}(kn) - S_{T_2}(kn-k)$$

$$= S_{T_1}(0) + \sum_{i=1}^{kn} G_{T_1}(i) - S_{T_1}(0) - \sum_{i=1}^{(k-1)n} G_{T_1}(i)$$

$$= \sum_{i=0}^{k-1} G_{T_1}(kn-i),$$

where $G_{T}(n)$ is the log-return obtained from the log-price $S_{T}(n)$ that is sampled with period $T$. Note that if $G_{T_2}(n) \sim N(\mu, \sigma^2)$, then $G_{T_2}(n) \sim N(k\mu, k\sigma^2)$. Using (10) and assuming that the mean of log-returns is zero, i.e. $\mu = 0$, we rewrite (3) as a function of the sampling period as

$$\rho_T = \frac{E[G_{T_1}(n)G_{T_2}(n)]}{\sigma_{T_1}\sigma_{T_2}},$$

(12)

where we assume that $E[G_{T_1}(n-k)G_{T_2}(n-l)] = \rho_T \sigma_{T_1}\sigma_{T_2}\delta_{k-l}$, i.e. the cross-correlation between different time periods of different asset log-returns is zero. Given (9), (12), and $\sigma_{T_2} = \sqrt{k}\sigma_{T_1}$, we conclude that

$$\rho_{T_1} = \rho_T,$$

(13)

It is shown that the cross-correlation coefficient, $\rho$, between the returns of two assets that follow geometric Brownian motion paths with pure Gaussian increments is not related to the sampling period. However, Epps [6] was the first to show that the empirical data does not comply with (13). Moreover, Epps stated that the cross-correlation between two financial assets decreases as the sampling period gets smaller, i.e.

$$\rho_T \to 0 \text{ as } T \to 0.$$

Cross-correlation coefficient between the log-returns of Apple Inc. stock (AAPL) and PowerShares QQQ Trust ETF (QQQ) as a function of time is displayed in Fig. 1.a. The sample correlation estimator employed for Fig. 1 is written as

$$\hat{\rho}_T = \frac{1}{W} \sum_{i=0}^{W-1} G_{T_1}(n-i)G_{T_2}(n-i),$$

(14)

where $G_T(n)$ is the normalized log-return with zero mean and unit variance as calculated

$$\hat{G}_T(n) = \frac{[G_T(n) - \hat{\mu}_T]}{\hat{\sigma}_T}$$

$$\hat{\mu}_T = \frac{1}{W} \sum_{i=0}^{W-1} \hat{G}_T(n-i)$$

$$\hat{\sigma}_T^2 = \frac{1}{W} \sum_{i=0}^{W-1} [\hat{G}_T(n-i) - \hat{\mu}_T]^2,$$

and $W$ is the measurement window size. Since AAPL is an important member of NASDAQ100 index and QQQ mimics the behavior of the index, one expects to have a significant correlation between the returns of these two relevant assets.
However, Fig. 1.a suggests a more complicated information. It is observed from the figure that assumption of (13) does not always hold, and the Black-Scholes model for the log-price of the assets (8) needs to be improved. This concern has been an active research topic where several people forwarded improved models (see [3, 4] and references therein.) Some researches have also proposed improved versions of estimators given in (14) and (15) [5].

The most widely accepted cause of the Epps effect in finance is the nature of asynchronous trading. Although prices of the assets within the same industry tend to behave similarly, and they respond to various intra-day economical and socio-political news in the same way, they are not traded at the same time points. The built-in asynchronicity affects correlation calculation since even the price of one asset of the pair does not change, the product of returns becomes zero, and that term is considered in the averaging operator. Similarly, a situation where the constant prices of both assets traded synchronously are also considered as perfectly uncorrelated pair of returns within this framework. Although these two zero-product cases for correlation calculations are distinct, it is more likely not to have price change at smaller time intervals. It is reasonable to expect higher correlation if correlation calculation only considers nonzero products with irregular sampling grid where price variations occur for both assets. Moreover, even though they were traded at the same time, their market structures, i.e. order books, volume and liquidity, are different and they

Figure 1. (a) Cross-correlation between AAPL and QQQ stocks estimated using 60 days of historical data between Apr 1, 2010 and Jun 30, 2010 as a function of sampling period, $T$. (b) A typical snapshot of the first five levels of the orderbooks (on both bid and ask sides) for AAPL and QQQ. Normalized last traded prices of both stocks on Mar 17, 2011 (c) between 9:30am and 9:32am sampled with $T = 1s$, (d) between 9:30am and 4:00pm sampled with $T = 300s$. 


and second largest five holdings are displayed in 2.a and 2.b, respectively, as a function of sampling periods. It is observed from these figures that the correlation pattern and Epps Effect observed in 1.a. are not specific to AAPL and QQQ pair. The pairwise stock correlations are widely utilized in risk management methods and electronic trading algorithms. Therefore, their calculation and proper use in financial systems is of crucial importance.

Fig. 3.a displays histogram of pairwise product operations in the correlation calculation of AAPL and QQQ pair at 1s sampling period. Similarly, Fig. 3.b displays histogram for the case of EOD correlation. Fig. 3.c depicts probabilities of product terms being negligible (between $-\epsilon$ and $\epsilon$), and used in the correlation calculations as a function of sampling period for $\epsilon = 3 \cdot 10^{-6}$. It is noted that the probability of having negligible product term in correlation calculation drops when sampling (trading) interval increases as depicted in this figure. This fact has direct impact on the values of pairwise correlations calculated through averaging that includes negligible ones. Hence, they drop significantly at higher sampling frequencies. This point needs further study.

IV. CONCLUSIONS

In this paper, we revisited Epps effect and pairwise correlation calculations of asset returns. The latter is a widely used metric in finance and plays a fundamental role in portfolio optimization employing Markowitz’s MPT, in hedging, and pairs trading. We offered several insights with supporting evidence to explain Epps effect. We expect further improvements in modeling of pairwise correlations for assets in an investment portfolio in order to develop more sophisticated hedging methods in the future.

REFERENCES