amadeus capital

Click or scan the QR code to visit us online

Estimating correlations across time zones

Fabian Scheler, CFA, FDP[1,2]

[1]*Amadeus Capital SA* [2]*Amadeus Quantamental SARL*

- Investors managers use correlation matrices to construct optimal portfolios.
- Periodicity mismatches induced by time-shift can severely distort the correlation estimates obtained for assets traded in different time zones.
- We empirically test the difference of correlation coefficients obtained over different rolling periods.
- Investors should pay attention to how they derive correlation matrices to avoid overestimating diversification benefits.

March 27, 2020

Since its introduction by Harry Markowitz in 1952, Modern Portfolio Theory has surely become one of the most influential concepts in Finance. The notion that "diversification is the only free lunch in investing." has shaped the thinking of academics and investment professionals around the world. Despite the concept's great recognition, though, according to our experience, investors tend to apply it in a rather general, less formalized way.

1 Correlation matrices in portfolio construction

The fallacies of classical quantitative portfolio optimization have been the topic of numerous research papers. Especially the difficulty of estimating expected returns and the finding that optimizers tend to suggest corner portfolios have been widely discussed. In this context, we highly appreciate the Black-Litterman model as a practical optimization problem solution. The Black-Litterman approach mitigates both previously mentioned fallacies by using implied expected returns derived from the market portfolio. It thereby avoids the difficulty of estimating asset returns and yields well-diversified portfolios. Furthermore, it allows active investors to efficiently embed the results of their research, thereby acknowledging the potential of positive contribution from human skills.

Nevertheless, just like the classical Markowitz optimization, the Black- Litterman model relies on the covariance matrix as a critical input parameter. Furthermore, even less quantitatively oriented investors take allocation decisions based on correlation analysis. Therefore, this article addresses one of the major fallacies in estimating covariance matrices - the comparison of time series across different time zones. We show the magnitude of this effect by the example of major equity indices and illustrate one possible solution for practitioners.

2 The problem with non-continuous trading

Investors who work with time series are used to the problem that historical data is usually limited to the open, high, low, and close prices. We are constrained by one data point per day or even less for many assets such as total return indices or investment funds. This becomes a significant problem when we compare time-series across different regions where the closing

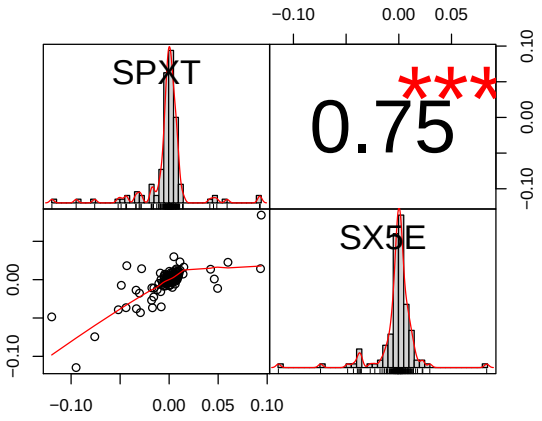

Figure 1: *Naïve Correlation Analysis*

price of one market already contains information that was not yet available when the other market closed. Let's take two of the most widely followed equity indices, the S&P500 and the EuroStoxx50, as an example. The graphic on the left shows the correlation between the daily returns of the S&P500 the EuroStoxx50 total return indices between September 2019 and today. Returns are calculated as the percentage difference between the closing prices of each index on each trading day and the preceding trading day.

Non-congruent public holidays are ignored by setting the return of the respective index on a respective day to zero. As expected, the correlation between both indices is pretty high and statistically significant at the 0.1% level. Squaring the correlation yields R2 of 0.56, meaning approximately 56% of the daily movements of the EuroStoxx50 can be explained by the behavior of the S&P500 and vice versa. This, however, does not account for the fact that based on Central European Time, trading of the EuroStoxx50 ends at 17:30, while the S&P500 closes 3.5 hours later at 21:00. To estimate

DateTime	SPX	SX5E	TPX	UKX
2019-09-11 09:00:00	3513	*NA*	*NA*	7313.5
2019-09-11 09:30:00	3518	*NA*	1586.0	7315.0
2019-09-11 10:00:00	3512	*NA*	1582.5	7334.0
2019-09-11 10:30:00	3514	2980.9	1584.2	7343.0
2019-09-11 11:00:00	3507	*NA*	1582.5	7328.5
2019-09-11 11:30:00	3511	*NA*	1583.5	7330.5
2019-09-11 12:00:00	3514	2979.5	1584.0	7338.0
2019-09-11 12:30:00	3512	*NA*	1582.0	7340.5
2019-09-11 13:00:00	3514	*NA*	1582.7	7338.5
2019-09-11 13:30:00	3518	2979.8	1584.0	7342.5

Figure 2: *Continuous Futures Data Snapshot*

the effect of this non-congruency in trading activity on the correlation estimate, we therefore calculate returns that span over the same time periods. We do so not only for the S&P500 but also for the British FTSE100 and the Japanese TOPIX. We, therefore, resort to the continuously traded Futures and retrieved 30 minutes bars from Bloomberg. This intra-day data is only available for the past six months which is why we chose the period from September 2019 to now. The table on the left shows a snapshot of the data. As can be seen, there is still a problem with the lack of trading activity for the EuroStoxx50 and the TOPIX.

To minimize the effect of this, we remove all rows where data is missing for one of the futures, including public holidays. We then select the latest point in time on each day for which all four futures show a price to calculate daily returns. This usually turns out to be 21:30 Central European Time. The following chart shows the correlations between daily returns for all four futures after adjusting for the different time zones. As can be seen, the correlation between the S&P500 and the EuroStoxx50 increases from 75% to 94%. Compared to this, the correlation between the EuroStoxx50 and the TOPIX, as well as the TOPIX and the UKX, still seem relatively low, though. As previously

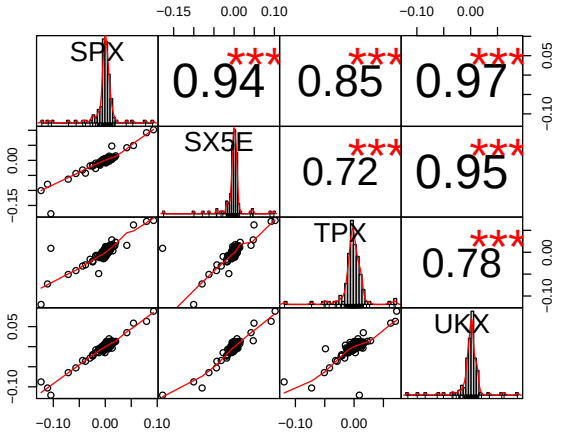

Figure 3: *Adjusted Correlation Analysis*

shown, we face a problem with missing data due to longer breaks with no trading activities in both indices. Therefore, in a second step, I form pairwise subsets of the four-time series and estimate correlation based on optimal return intervals for the respective pair. Results are mostly the same but correlation between the TOPIX and the FTSE100 increases from 0.78 to 0.9. As outlined before, the availability of intraday data is generally limited. However, Bloomberg provides ticks and bars to ordinary subscribers for up to six months. The download of more extended time-series requires a costly data license. Furthermore, intra-day prices are generally only available for some assets.

Noncontinuous trading activity in most markets and limited overlap in trading hours between different timezones provide another challenge. The exercise above has shown that daily correlations between significant equity markets are way higher than a simple end-of-day-based analysis would suggest. Investors who base allocation decisions and risk management on covariance matrices derived without further adjustments are likely to underestimate their exposure to systematic risk and overestimate diversification. However, estimating better parameters based on intra-day data is not feasible in most cases.

We, therefore, suggest another method, namely the estimation of the covariance matrix using contemporary and lagged returns. Let's revisit the first dataset containing the end of day prices of the total return indices for the S&P500 and the EuroStoxx50.
We first run a multiple regression of the EuroStoxx50 returns on the return of the S&P500 on the concur-

	Dependent variable:
	SX5E
SPXT	0.792***
	(0.042)
SPXT_lag1	0.444***
	(0.048)
SPXT_lag2	-0.001
	(0.047)
SPXT_lag3	-0.128**
	(0.047)
Constant	-0.00003
	(0.001)
Observations	135
R^2	0.743
Adjusted R^2	0.735
Residual Std. Error	0.009 (df = 130)
F Statistic	93.986*** (df = 4; 130)
Note:	*p<0.1; **p<0.05; ***p<0.01

rent trading day and the previous trading days. We thereby identify the loading of EuroStoxx50 returns on the lagged returns of the S&P500. Results are presented in the table on the left. As expected, due to the difference in closing time, the EuroStoxx50 shows a highly significant, positive loading on the returns of the S&P500 on the previous trading day. This is also relevant when estimating beta factors. The simple regression yields a securities beta of only 0.79. However, regression betas are additive, which means that taking the statistically significant loadings from the multiple regression, the EuroStoxx50 has a daily beta of 1.236 against the S&P500. Most importantly, the regression's R2 is 0.74, implying a correlation of 0.86 compared to the R2 of 0.56 in the previous correlation analysis.

Unfortunately, we are not aware of any major financial data vendor, providing security betas estimated through multiple regressions using contemporary and lagged returns. Based on previous work on single stocks, we, however, suggest using the technique not only when dealing with different timezones but also when estimating risk factors for single stocks, especially when trading activity is low, and prices most likely don't react to market-wide news in a timely way. As dealing with multiple pairwise regressions becomes rather tedious when working with various assets, we also test a second, more straightforward approach and estimate correlation over rolling multi-day windows. The following charts show the correlation between the S&P500 and the EuroStoxx50 for rolling returns over windows with different lengths. In this case, adding just one more day does not change a lot, increasing the correlation from 0.75 to 0.76

only. However, choosing a longer window seems to be a highly effective way to mitigate the problem of non-contemporaneous trading activity. In this context, for risk management purposes, it seems to make sense to also test different window lengths to see the impact and thereby "stress-test" the covariance matrix. I repeat the same exercise for the EuroStoxx50 and the TOPIX and the EuroStoxx50 and the FTSE100. Results are shown below. The results for the correlation between EuroStoxx50 and TOPIX mirror those for the EuroStoxx50 and the S&P500. On the other hand, adding lag factors has only a minimal impact on the estimate of the correlation between the British FTSE100 and the EuroStoxx50. This makes sense as the difference in trading hours between both places is minimal. It, therefore, indicates that using rolling multi-day window returns does indeed capture the effect of non-synchronous trading activity.

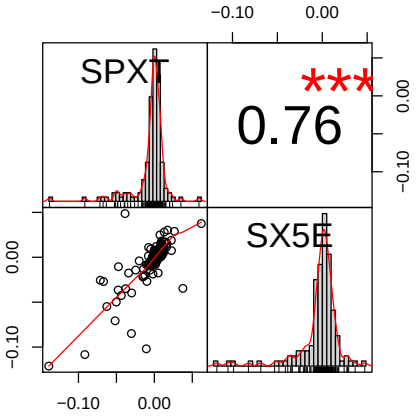

Figure 4: *Correlation Analysis 2 Days Rolling Returns*

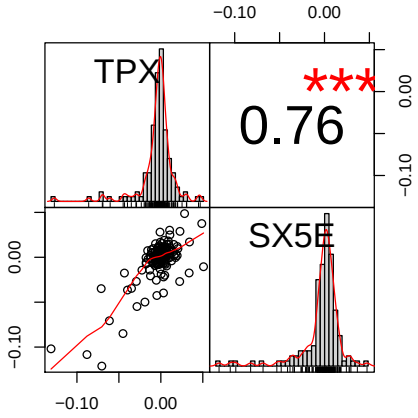

Figure 7: *Correlation Analysis 2 Days Rolling Returns*

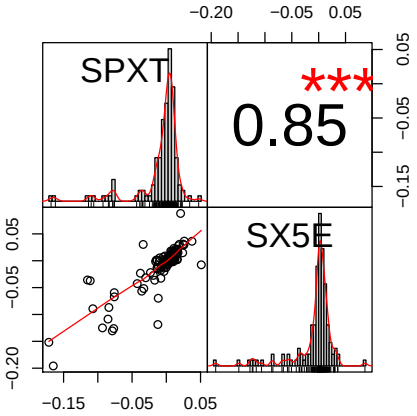

Figure 5: *Correlation Analysis 4 Days Rolling Returns*

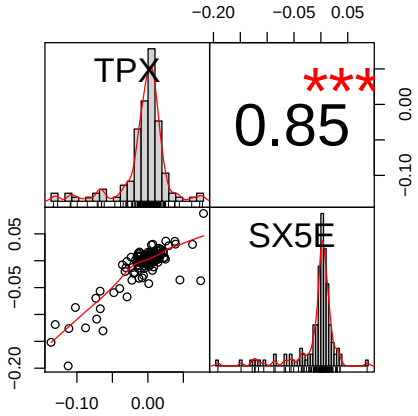

Figure 8: *Correlation Analysis 4 Days Rolling Returns*

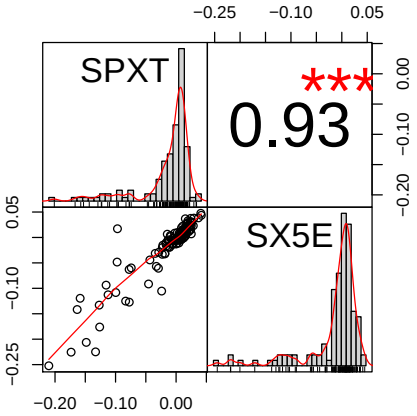

Figure 6: *Correlation Analysis 7 Days Rolling Returns*

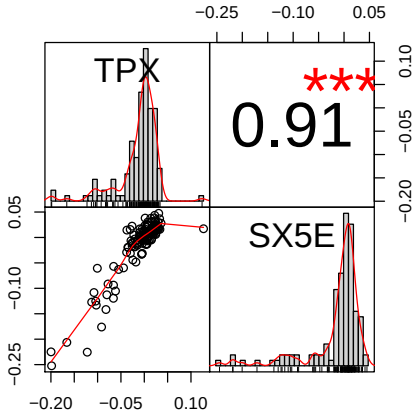

Figure 9: *Correlation Analysis 7 Days Rolling Returns*

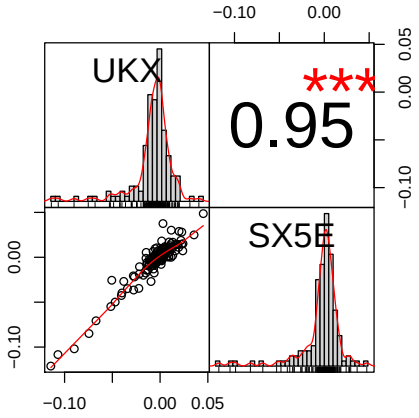

Figure 10: *Correlation Analysis 2 Days Rolling Returns*

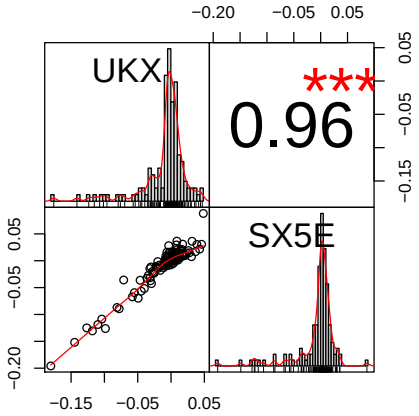

Figure 11: *Correlation Analysis 4 Days Rolling Returns*

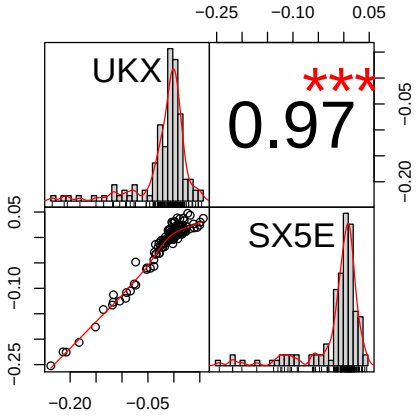

Figure 12: *Correlation Analysis 7 Days Rolling Returns*