SPECTRAL ANALYSIS OF STOCK-RETURN VOLATILITY, CORRELATION, AND BETA

A. Shomesh E. Chaudhuri

Massachusetts Institute of Technology
Electrical Engineering and Computer Science
Cambridge, MA

*B. Andrew W. Lo**

Massachusetts Institute of Technology
Sloan School of Management, CSAIL, EECS
Cambridge, MA

ABSTRACT

We apply spectral techniques to analyze the volatility and correlation of U.S. common-stock returns across multiple time horizons at the aggregate-market and individual-firm level. Using the cross-periodogram to construct frequency band-limited measures of variance, correlation and beta, we find that volatilities and correlations change not only in magnitude over time, but also in frequency. Factors that may be responsible for these trends are proposed and their implications for portfolio construction are explored.

Index Terms— spectral analysis, volatility, correlation, beta, portfolio theory, financial engineering

1. INTRODUCTION

It has been observed that the volatility of securities and their correlation with the stock market are not constant, but change over time. This variation in volatility and correlation has important implications for any theory of risk, return, and portfolio construction. However, standard measures do not distinguish between the short- and long-term components of risk and co-movement. The fact that economic shocks produce distinct effects on stock return dynamics at different time horizons suggests that frequency-specific measures of volatility and correlation may yield several new insights.

First, studying the frequency components of stock return processes can reveal new features of the underlying economic structure driving these processes. Second, since the time horizon for investments can range from microseconds to decades, the risks specific to these time horizons can be gauged and taken into consideration during portfolio construction. Third, as trading strategies become ever faster, the frequency domain provides a convenient framework for comparing investment processes that operate on different timescales and, more importantly, for diversifying across these timescales. Finally, frequency-domain measures offer surprisingly simple representations of complex dynamics that are cyclical.

In this article, we propose several new frequency-domain measures of stock-market risk and expected return and show how they can complement traditional approaches to portfolio management. We begin in Section 2 by reviewing the literature on volatility, portfolio theory, and spectral analysis as applied to finance. In Section 3 we present the basic decomposition of variance, correlation, and beta into their frequency components, and in Section 4 we apply this framework to the historical returns of U.S. stocks. We explore the applications of these measures to portfolio theory in Section 5, and conclude in Section 6.

2. LITERATURE REVIEW

There is an enormous literature on statistical models for capturing the time variation in volatility. Simple models such as the rolling standard deviation used by Officer [1] have given way to sophisticated econometric techniques such as ARCH and GARCH models introduced by Engle [2] and other stochastic volatility models. Partial surveys of these methods are given by Campbell, Lo, and MacKinlay [3], Broto and Ruiz [4], and Bauwens, Laurent and Rombouts [5]. In this article, we focus on characterizing the time-horizon-dependent components of volatility and correlation; hence, our use of spectral analysis.

Spectral analysis has a long history in econometrics [6], [7], with applications ranging from business cycle analysis [8] to option valuation [9]. Often the analysis relies on the Fourier transform (e.g., [10], [11]), however recently wavelets (e.g., [12], [13], [14]) and the Hilbert-Huang transform (e.g., [15]) have also been used to study financial data in the time-frequency domain. In this article, the cross-periodogram—calculated using the short-time Fourier transform [16]—forms the basis of our spectral analysis. The cross-periodogram, which is the decomposition of the inner product of two time series into their frequency components, was introduced into the economic literature by Engle [17] as a component of band-spectrum regression. Using a similar insight, we use the frequency-band-limited adaptations

*The views and opinions expressed in this article are those of the authors only and do not necessarily represent the views and opinions of any other organizations, any of their affiliates or employees, or any of the individuals acknowledged above. Research support from the MIT Laboratory for Financial Engineering is gratefully acknowledged.

978-1-4673-9169-6/15/$31.00 ©2015 IEEE

of variance, correlation, and beta, to study the statistical properties of asset returns at multiple time horizons. This framework connects spectral analysis to the standard tools of modern portfolio theory developed by Markowitz [18], allowing us to analyze the time-horizon properties of portfolio construction.

3. SPECTRAL ANALYSIS

We use the discrete Fourier transform (DFT) to decompose the covariance of returns into its frequency components, from which we calculate time-horizon-specific correlations and betas.

Let $r_{i,t}$ be the one-period return of stock i between dates $t{-}1$ and t. The sample variance of returns over a interval from $t = 0, \ldots, T{-}1$ can then be calculated as:

$$\mathrm{var}(r_i) \;=\; \frac{1}{T}\sum_{t=0}^{T-1}(r_{i,t} - \overline{r}_i)^2 \,, \tag{1}$$

where $\overline{r}_i$ is the sample mean of returns,

$$\overline{r}_i \;=\; \frac{1}{T}\sum_{t=0}^{T-1} r_{i,t} \,. \tag{2}$$

This calculation is exactly equivalent to the one formed by averaging over the periodogram:

$$\mathrm{var}(r_i) \;=\; \frac{1}{T^2}\sum_{k=1}^{T-1} R_{i,k}R^*_{i,k} \tag{3}$$

where $R_{i,k}$ are the T-point DFT coefficients[1] of $r_{i,t}$:

$$R_{i,k} \;=\; \sum_{t=0}^{T-1} r_{i,t}e^{-j\frac{2\pi k t}{T}} \quad , \quad 0 \le k \le T{-}1 \,. \tag{4}$$

Note that $k = 0$, the zero frequency, is not involved in (3) since adding or subtracting a constant to the time series $r_{i,t}$ does not change its sample variance. In this form, the contribution to the sample variance by the kth frequency, where $k \in \{1, \ldots, T-1\}$, is real valued and clearly visible. The lowest non-zero frequency occurs at $k = 1$ and the highest frequency occurs at the value of k closest to $T/2$. Values of k that are symmetric about $T/2$ (e.g., $k = 1$ and $k = T-1$) have the same frequency and their contributions to the sample variance are the same. The relation $h = TT_s/k$, where T_s is the time between samples and $0 \le k \le T/2$, can be used to convert the kth frequency to its corresponding time horizon.

[1] These coefficients can easily be calculated using the algorithms provided by any standard statistical software program. For example, we used the "`fft`" function in MATLAB.

More generally, averaging over the cross-periodogram can be used decompose the sample covariance into its frequency components:

$$\mathrm{cov}(r_i, r_j) \;=\; \frac{1}{T^2}\sum_{k=1}^{T-1} R_{i,k}R^*_{j,k} \,. \tag{5}$$

This technique uses the DFT to express the returns in the frequency domain and then analyzes their phase. When the returns are in phase at a given frequency, the contribution that frequency makes to the overall covariance is positive. When they are out of phase, then that particular frequency's contribution will be negative. Values of k that are symmetric about $T/2$ (e.g., $k = 1$ and $k = T-1$) have the same frequency and their contributions to the sample variance are complex conjugates. Therefore, the contribution from a single frequency or subset of frequencies will always be real.

We can calculate the contribution of a subset of frequencies, $K \subseteq \{0, \ldots, T{-}1\}$, to the sample covariance by limiting the summation in (5) to the frequencies of interest:[2],

$$\mathrm{cov}_K(r_i, r_j) \;=\; \frac{1}{T^2}\sum_{k\in K} R_{i,k}R^*_{j,k} \,. \tag{6}$$

This corresponds to calculating the sample covariance of the inverse DFT reconstruction of returns, restricted to the specified frequency subset.

A few important implementation details still remain. Windowing procedures (e.g., multiplication by a Hamming window) can be applied to the data before taking the Fourier transform used to calculate the modified periodogram. This procedure will generally decrease spectral leakage at the expense of reducing spectral resolution. Moreover, averaging cross-periodograms formed over overlapping time intervals can reduce the variance of the spectral estimates at the expense of increased bias. This method is known as the Bartlett method if no window other than a rectangular window is applied to the data sections, and as Welch's method if a tapered window is applied. We refer the reader to Oppenheim and Schafer [16, Chapter 10] for a more detailed discussion of these advanced implementation techniques.

This frequency-domain analysis can be extended to time-horizon specific calculations of the beta and correlation coefficients. The frequency-restricted beta coefficient, β_K, and correlation coefficient, ρ_K, can be calculated as:

$$\beta_K(r_i, r_j) \;=\; \rho_K(r_i, r_j)\left[\frac{\mathrm{var}_K(r_i)}{\mathrm{var}_K(r_j)}\right]^{1/2} \tag{7}$$

where

$$\rho_K(r_i, r_m) \;=\; \frac{\mathrm{cov}_K(r_i, r_j)}{[\mathrm{var}_K(r_i)\mathrm{var}_K(r_j)]^{1/2}} \,. \tag{8}$$

[2] Note that pairs of elements that correspond to the same frequency should be included together in K.

Since returns are real-valued, $-1 \leq \rho_K(r_i, r_j) \leq 1$, and (8) is computationally equivalent to calculating the sample correlation of the inverse DFT reconstructions of returns, restricted to the frequencies specified by K. Engle [17] shows how the band spectrum regression coefficient β_K can be generalized to the case of multiple factors.

4. MEASURING TRENDS IN VOLATILITY AND CORRELATION

Our empirical analysis focuses on all stocks in the University of Chicago's CRSP Database from May 10, 1963 to December 31, 2014. Specifically, we use only U.S. common stocks with CRSP share code 10 and 11, which eliminates REIT's, ADR's, and other non-common-stock securities. CRSP occasionally reports returns and prices that are based on bid-ask quote midpoints (for infrequently traded securities); we eliminate a stock from our sample if more than 5% of its prices are based on these values in any windowed time interval.

In Fig. 1 we show the trailing 750-trading-day annualized variance and variance spectrum of the CRSP value-weighted market index. We compute the annualized variance based on daily data. The figure shows large spikes in volatility across all frequencies during the oil shocks of the 1970s, the stock market crash of 1987, the Tech Bubble, and the Financial Crisis of 2007–2008. Moreover, the variance spectra over the past two decades show more high-frequency content than any of the earlier periods, which is not surprising considering the ever increasing speed and volume of trading.

Figs. 2 and 3 show the cross-sectional average of the trailing 750-trading-day annualized variance and correlation spectra when applied to stocks sorted into NASDAQ and NYSE/AMEX subsets by their CRSP exchange code identifier. We find that, throughout the 1990s, the average variance of NASDAQ securities trended upwards, especially at the shorter time horizons (between 1 day and 1 week). In contrast, there is no discernible trend in the average variance of stocks trading on the NYSE/AMEX during this period. Since this upward trend in short-run individual stock volatility did not translate into increased market volatility, we see in Fig. 2 that, as expected, the average correlation at these high frequencies decrease. We also note that the correlation spectra in both Figs. 2 and 3 show large increases in the average correlation with the market across all frequencies during the stock market crash of 1987 and the Financial Crisis of 2007–2008. Moreover, correlations with the market have tended to trend upward over the past two decades.

It is natural to ask what might explain the increasing volatility and decreasing market correlation observed in these NASDAQ securities during the 1990s. We speculate that this trend may be partly explained by the high-frequency dynamics induced by the advent and proliferation of electronic trading networks during this period. NASDAQ embraced electronic trading much earlier than the NYSE/AMEX and

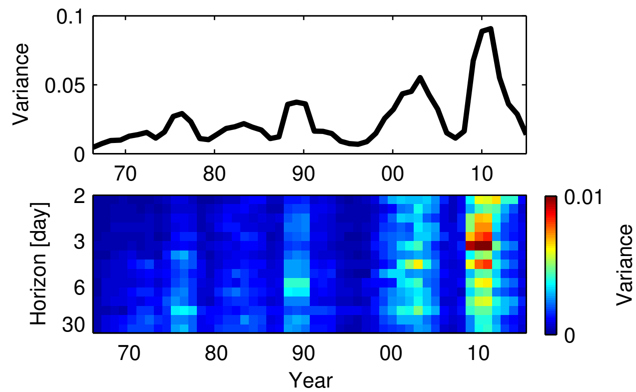

Fig. 1. The top panel shows the trailing 750-trading-day annualized variance of the value-weighted market index provided in the CRSP data set between May 10, 1963 and December 31, 2014. The bottom panel shows the corresponding variance spectrum partitioned into 15 frequency bands.

provided various electronic platforms for both market makers and customers. The algorithmic trading strategies employed on these electronic platforms may have increased the amount of noise at these shorter time horizons, while remaining uncorrelated with general market movements. These trends would have been particularly noticeable in stocks with low baseline prices and small market capitalizations—such as those found predominantly on the NASDAQ stock exchange—since high-frequency price dynamics on the order of the bid-ask spread would have subsequently larger effects on returns.

An interesting area for further research is to determine the underlying processes driving these volatility and correlation measures, and, in particular, explaining the upward trend in high-frequency volatility in the 1990s.

5. IMPLICATIONS FOR PORTFOLIO CONSTRUCTION

We have shown that individual stock volatility and correlations change over time and across frequencies. An implication of this finding is that a portfolio can be designed to maximize expected returns while minimizing its risk over a given frequency band. Since investors have different time horizons, it is rational for them to seek to minimize the risk specific to a particular frequency band. For example, short-run volatility, even if correlated with their portfolio, may not be important to their investment goals if their time horizon is much longer. Similarly, low-frequency power in returns and correlations may be insignificant to a high-frequency trader who does not operate on the same timescale. Our frequency decompositions provide a systematic framework for performing such an analysis.

In mean-variance portfolio theory, given a target value

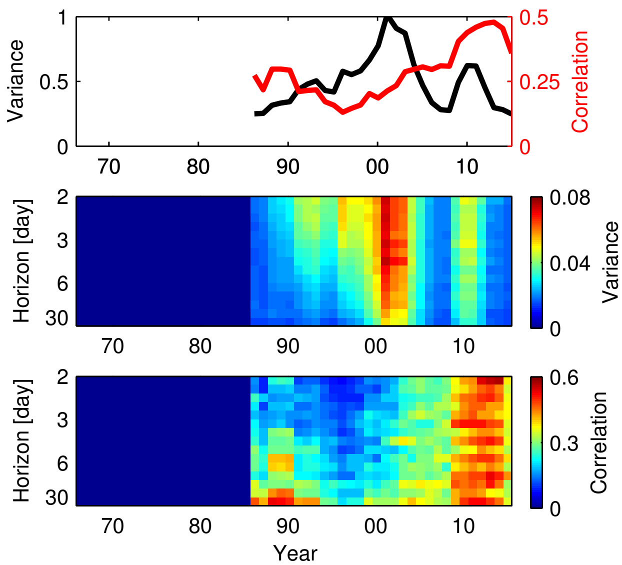

Fig. 2. The top panel shows the cross-sectional average for NASDAQ securities of their trailing 750-trading-day return annualized variance and correlation with the value-weighted market index return between April 13, 1983 and December 31, 2014. The middle and bottom panels show the corresponding average annualized variance and correlation spectra, respectively, partitioned into 15 frequency bands.

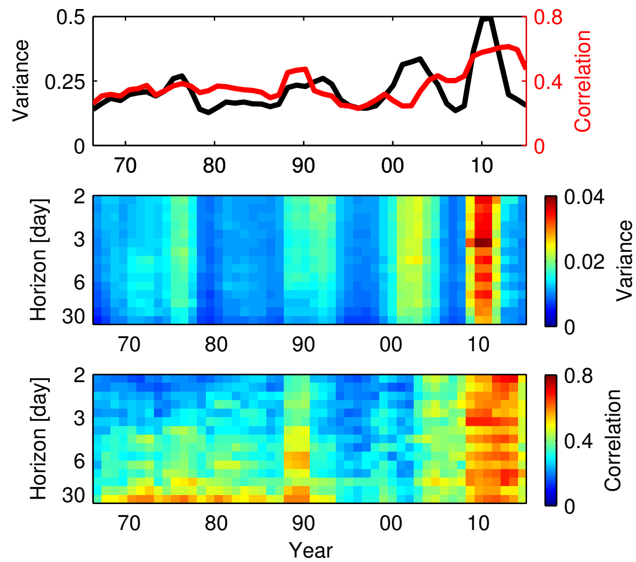

Fig. 3. The top panel shows the cross-sectional average for NYSE/AMEX securities of their trailing 750-trading-day return annualized variance and correlation with the value-weighted market index return between May 10, 1963 and December 31, 2014. The middle and bottom panels show the corresponding average annualized variance and correlation spectra, respectively, partitioned into 15 frequency bands.

($\tilde{\mu}$) for the expected portfolio return, the efficient portfolio weights ($\tilde{\mathbf{w}}$) are those that minimize the portfolio variance for all portfolios with expected return $\tilde{\mu}$. Mathematically, the optimization problem can be written as,

$$\tilde{\mathbf{w}} \quad = \quad \arg\min_{\mathbf{w}} \mathbf{w}^{\mathrm{T}} \boldsymbol{\Sigma} \mathbf{w} \tag{9}$$

subject to the constraints

$$\mathbf{w}^{\mathrm{T}} \boldsymbol{\mu} = \tilde{\mu} \text{ and } \mathbf{w}^{\mathrm{T}} \mathbf{1} = 1 \tag{10}$$

where w_i is the portfolio weight on the ith security, $\mu_i = \mathrm{E}[r_i]$ and $\Sigma_{i,j} = \mathrm{cov}(r_i, r_j)$.

The inputs for the optimization problem are, therefore, the expected returns and covariance matrix of these securities. As a first-order approximation, sample estimates can be used for these values. Accordingly, a time-horizon-specific mean-variance optimization, restricted to the frequency band K, can be developed by simply replacing the sample standard deviation and correlation matrix estimates with those based on (6) and (8), respectively. This framework has the attractive feature that the optimization techniques developed to solve for the efficient frontier are still valid since the form of inputs are not affected.

Now suppose that an investor currently holds portfolio A, and wishes to shift a small fraction of his portfolio weight, δ, to security i to create portfolio B. It can be shown that the ratio of the variance of portfolio B to that of portfolio A is approximately:

$$\frac{\mathrm{var}(r_B)}{\mathrm{var}(r_A)} \quad \approx \quad 1 \; + \; 2\delta(\beta(r_i, r_A) - 1) \tag{11}$$

which implies that if a small amount of security i is added to portfolio A, the variance of the portfolio will increase if $\beta(r_i, r_A) > 1$, and will decrease if $\beta(r_i, r_A) < 1$. Taking advantage of the linear properties of the Fourier transform, we find that:

$$\frac{\mathrm{var}_K(r_B)}{\mathrm{var}_K(r_A)} \quad \approx \quad 1 \; + \; 2\delta(\beta_K(r_i, r_A) - 1) \tag{12}$$

which provides a similarly intuitive interpretation, in terms of portfolio construction, for the restricted-frequency beta coefficient, β_K.

These simple examples highlight the versatility of cov_K, ρ_K and β_K. As the frequency band-limited counterparts to covariance, correlation, and beta, their applications are not limited to modern portfolio theory, but can be applied to almost any theory of risk, reward, and portfolio construction. An interesting area for future research is to investigate the the practical advantages of such a framework.

6. CONCLUSION

In this article, we apply spectral analysis to characterize the behavior of stock-return volatility, correlation, and beta across multiple time horizons at the individual-firm and aggregate-market level. Our approach exploits a frequency-band analysis of variance and correlation to analyze the daily returns of securities trading on the NASDAQ and NYSE/AMEX stock exchanges. We identified periods of increasing variance and decreasing correlation in NASDAQ securities at the shorter time horizons between 1 day and 1 week. We speculate that this trend may result in part from changes induced by electronic trading. Finally, we suggest some potential applications for the frequency band-limited adaptations of covariance, correlation, and beta to portfolio theory. These considerations are particularly useful when portfolio goals differ across time horizons, and when investors wish to target specific horizons because of their preferences and life cycle.

7. REFERENCES

[1] R.R. Officer, "The variability of the market factor of the New York Stock Exchange," *Journal of Business*, vol. 46, pp. 434–453, July 1973.

[2] R.F. Engle, "Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation," *Econometrica*, vol. 50, pp. 987–1007, July 1982.

[3] J.Y. Campbell, A.W. Lo, and A.C. MacKinlay, *The Econometrics of Financial Markets*, Princeton University Press, 1997.

[4] C. Broto and E. Ruiz, "Estimation methods for stochastic volatility models: A survey," *Journal of Economic Surveys*, vol. 18, pp. 613–649, December 2004.

[5] L. Bauwens, S. Laurent, and J.V.K. Rombouts, "Multivariate GARCH models: a survey," *Journal of Applied Econometrics*, vol. 21, pp. 79–109, 2006.

[6] C.W.J. Granger and M. Hatanaka, *Spectral Analysis of Economic Time Series*, Princeton University Press, 1964.

[7] C.W.J. Granger and R.F. Engle, "Applications of spectral analysis in econometrics," *Handbook of Statistics*, vol. 3, pp. 93–109, 1983.

[8] M. Baxter and R.G. King, "Measuring business cycles: Approximate band-pass filters for economic time series," *review of economics and statistics*, vol. 81, pp. 575–593, November 1999.

[9] P. Carr and D. Madan, "Option valuation using the fast fourier transform," *Journal of computational finance*, vol. 2, pp. 61–73, 1999.

[10] C. Croux, M. Forni, and L. Reichlin, "A measure of comovement for economic variables: Theory and empirics," *Review of Economics and Statistics*, vol. 83, pp. 232–241, May 2001.

[11] J. Breitung and B. Candelon, "Testing for short- and long-run causality: A frequency-domain approach," *Journal of Econometrics*, vol. 132, pp. 363–378, June 2006.

[12] J.B. Ramsey, "Wavelets in economics and finance: Past and future," *Studies in Nonlinear Dynamics & Econometrics*, vol. 6, pp. 1–27, 2002.

[13] A. Rua, "Measuring comovement in the time-frequency space," *Journal of Macroeconomics*, vol. 32, pp. 685–691, June 2010.

[14] A. Rua, "Wavelets in economics," *Economic Bulletin and Financial Stability Report Articles*, pp. 71–79, 2012.

[15] N.E. Huang, M.L. Wu, W. Qu, S.R. Long, S.S.P. Shen, and J.E. Zhang, "Applications of Hilbert-Huang transform to non-stationary financial time series analysis," *Applied stochastic models in business and industry*, vol. 19, pp. 245–268, 2003.

[16] A.V. Oppenheim and R.W. Schafer, *Discrete-Time Signal Processing*, Prentice Hall, 3rd edition, 2009.

[17] R.F. Engle, "Band spectrum regression," *International Economic Review*, vol. 15, pp. 1–11, February 1974.

[18] H. Markowitz, "Portfolio selection," *The Journal of Finance*, vol. 7, pp. 77–91, March 1952.