**DRAFT**
**Financial Management** DOI: 10.1002/ISBN......

# A measure of portfolio informational quality based on the dynamics of equity co-movement

**William Smyth**[a]

**Keywords:** *Co-movement*;*Correlation*; *US Equity*; *Gerber Statistic*; *Smyth-Broby adjustment*;

**Abstract:** We introduce the concept of portfolio informational quality (portfolio-IQ) and propose a technique to measure it. We also demonstrate how portfolio-IQ can be used to improve portfolio optimisation outcomes. An axiomatic starting point is that market efficiency, as an absolute entity, is not uniquely measurable. However, we argue that informational inefficiency can be accounted for within the context of portfolio-IQ. The latter combines the effects of informational inefficiency and temporal changes in the underlying co-movement dynamics of portfolio components. By considering a bespoke portfolio of several asset classes, between January 1988 to December 2020, we demonstrate how our approach can identify a portfolio-specific informational quality resonance frequency. We use the Gerber statistic (incorporating a single modification) to explain our results, taking advantage of its robustness and superior noise elimination properties. By incorporating portfolio-IQ we extend the reach of the Gerber statistic in classifying co-movement data and, in doing so, add to the growing literature on the use of co-movement in portfolio management.

## 1. Introduction

In this paper we propose a technique which can be used to incorporate portfolio-specific informational quality into Markowitz optimisation. We take a phenomenological approach based on the assumption that the effects of market inefficiency can be evidenced through the co-movement dynamics of a combination of stocks comprising an index portfolio. This index portfolio is further assumed to be a proxy for a market (Levy & Roll (2010)). Motivated by actual rather than theoretical dynamics, this allows us to address how the collective of co-movement temporal properties and informational efficiency can be accounted for, thereby addressing a gap in the literature.

Need to edit all below in this section.......

It is accepted by finance scholars that, in theory, only common systemic components of return should generate excess return (Fama (1991). There is evidence, however, that in practice, and without new information, idiosyncratic movement in the prices of securities can be greater (or less) than justified by common factors of return (Longin &

[a]**Ulster University, Accounting, Finance and Economics Department, Cathedral District, Belfast BT15 1ED, United Kingdom. w.smyth@ulster.ac.uk**

 

Solnik (2001), Breymann et al. (2003), Bekaert et al. (2009)). Such un-priced risk, termed co-movement, presents a challenge to the Efficient Market Hypothesis (EMH) as only new information should be priced into stocks (Barberis et al. (2005)). If co-movement can not be explained and/or persists over time, finance theory would need to be revisited. Using co-movement measures, we therefore investigate the way that co-movement manifests itself over time. Our working hypothesis being that an age-related reduction in the informational value of observed co-movement is evidence of portfolio informational efficiency. Our results extend the EMH measurement tools and add to what is known about covariance and portfolio optimization by explaining how co-movement decay is a function of dynamic market efficiency.

Due to the computational issues involved with large correlation matrices, scholars have not researched co-movement in as much depth as it deserves, and tend to refer to it as statistical noise. As a result, equity co-movement and dependence have historically been investigated in bivariate and static contexts (Longin & Solnik (2001) and Poon et al. (2004)). Part of the issue with its investigation is that co-movement is extremely difficult to visualize, particularly in larger portfolios (Phoa (2013). In our own investigation, we focus exclusively on smaller portfolios. The other challenge for researchers is that the traditional approach to portfolio optimization is subject to a great deal of estimation error of the efficient frontier (Jobson & Korkie (1980)). Despite this, co-movement research is being revisited because it has been shown to contain useful information that can shape forward looking assumption (Chen et al. (2016)).

The reason multi-variate investigations into co-movement have proved challenging is due to the large number of variables involved (Aue et al. (2009)). Together with correlation, such movements are important to understand as they help determine the weights of optimal portfolios (Markowitz (1991). Two common co-movement measures, the Gerber statistic (GS) proposed by Gerber et al. (2022) and the modified Gerber statistic (SB) proposed by Smyth & Broby (2022) have been recently developed. These demonstrate superior risk reward portfolio outcomes compared to portfolios generated using the sample covariance matrix (or historical co-variance matrix) and those generated through shrinkage of the sample covariance matrix. In the light of the aforementioned challenges to finance theory, we use the aforementioned measures to address the following related questions: (i) Does the informational value of observed co-movement decay with age?, and (ii) Is there evidence that co-movement measures can deliver superior risk adjusted results by exploiting informational inefficiencies?

# 2. Framing Portfolio Informational Quality

text to introduce the next two subsections: Introduce this within the context of mean variance optimisation and the correlation matrix at the heart of this process. Estimating the correlation matrix is a challenging task. Empirical correlation matrix converges to the population covariance matrix as sample size increases. However this poses problems and analytical conflicts of interest. Specifically, on one hand between using samples (lookback windows) of sufficient length in order to gather a sample representative of the true dynamic versus using look back windows of such length that it is impossible to stand over the efficacy of older co-movement data. The latter is subject to:

1. unidentified regime change (a break in stationarity in the underlying returns data generating process 2. temporal changes in the co-movement relationships between assets [even in the absence of regime change] 3. market informational inefficiency.

Smyth  Broby (Quantitative Finance) proposed a stability measure technique to identify regime change allowing its impact to be factored into portfolio management. It is the latter two points we seek to address in this study. Each point is addressed in the following sections providing the basis for their collective consideration within the scope of a concept we define as portfolio informational quality. An idea central to our study.

Having provided a contextual framework to account for temporal changes in co-movement relationships and informational inefficiency we need a tool with which to implement these concepts in practice. In the Gerber statistic we have an ideal candidate for such a mission. The Gerber statistic, having been tried and tested in industry, formally entered the finance literature in 2022. As a robust statistical measure of co-movement, it facilitates building a covariance matrix using a technique which has been demonstrated to mitigate against noise better than other methods such as correlation and shrinkage. Clearly noise management is critically important in Markowitz optimisation due to the central role of the covariance matrix. We take advantage of this robustness as the basis for estimating the covariance matrix. We also take advantage of the structure of the Gerber statistic which is ideal in allowing us to incorporate portfolio informational quality analytically and numerically in a simple and transparent way.

## 2.1. Informational efficiency

put the stuff in here relating to market efficiency

## 2.2. Co-movement temporal dynamics

According to the EMH, proposed by Fama (1970), markets that provide accurate information signals reflect that in their pricing. The EMH is based on the concept of a "random walk" in securities prices, where successive share price changes are effectively random and detached from the previous price. Many anomalies and observed patterns have since been documented by academics (Malkiel (1989). This could be attributed to the conditions of the random walk theory, namely that price changes have to be both independent and conform to a probability distribution (Cheng & Deets (1971)). Addressing this assumption, LeRoy (1989) proposed a martingale approach be used to reflect the observed fact that securities experience varying levels of turbulence and volatility. In this way, the conditional variances of securities can be said to be positively auto-correlated.

Fama (1970) proposed three forms of the Efficient Market Hypothesis, (i) the weak, (ii) semi-strong and (iii) strong form. The weak-form is based on historical prices. The semi-strong form is based on historical prices but includes all publicly available information. The strong form is based on the former but includes non-public information. As per Roll (1984), we note that surprise co-movements can occur in informationally efficient markets. Several measures have been developed to capture degrees of efficiency in the market. These include those that measure whether there is a random walk using long run return predictability and/or variance (Darrat & Zhong (2000)). Also, those that measure whether there is a random walk using short run return predictability and/or variance using intraday data (Heston et al. (2010)). Our approach proposes a joint cross-sectional and longitudinal measure.

Long term decay has been observed in the nature of co-movement in the US (Campbell et al. (2001)). Our results are supportive of the semi-strong form of the efficient market hypothesis. That is, the conditional expected value of the future price, given all the past prices, is equal to the most recent price.

The validity of our claim that market inefficiency can be partially evidenced through the co-movement dynamics of a collection of stocks is demonstrated empirically. Robust co-movement statistics such as the Gerber statistic (GS) or the Smyth-Broby adjustment (SB) to the Gerber statistic are ideal candidates for this task. The Gerber family (a term we shall use to refer to GS and SB collectively) have superior noise elimination properties when compared to commonly used techniques of co-movement measurement such as correlation and shrinkage.

A bespoke feature of the Gerber family is that they individually classify each and every co-movement within a predetermined rolling window. Small magnitude co-movements are eliminated if they do not meet a certain threshold as it is impossible to be sure that a small magnitude co-movement is evidence of an underlying connection between two stocks or simply noise. Large magnitude co-movements have their impact moderated though the use of a one-size-fits-all count. Thus, the obfuscating effect of noise is mitigated for at both ends of the magnitude scale.

It should be noted that the Gerber family perform this mitigation in different ways but they share a common guiding principle. Any co-movement which passes the small magnitude threshold is referred to as a qualifying co-movement. All qualifying co-movements are categorised as being concordant, discordant or neutral. A running total is maintained of the counts in each category and collectively they evaluate a co-movement matrix element for a given asset-pair combination. Repeating this process for all possible asset-pair combinations populates the co-movement matrix. This highly structured and detail-focused approach provides an ideal framework within which to embed a market efficiency monitor.

In summary, we use the GS and SB co-movement measures to quantify the value-decay over time of co-movement information. Our resultant measure can be used to compare the cross sectional efficiency of different markets and even asset classes. Our findings are important because using our method, it is possible to say with a degree of confidence that one market is more efficient than another, something that finance academics have struggled with since the 1970's.

# 3. The literature on co-variance

There is substantial literature on the way securities vary and covary, as well as how they are correlated. Security returns, however, fail to reflect changes in covariances due to observed underlying information factors. Co-movement, for example, has been shown to be significantly higher during recessions (Veronesi (1999)). The United States equity market, as documented by Parsley & Popper (2020), has dominated much of the co-movement literature, largely because the longer term return data is widely available.

Academics have taken several different approaches to measure co-movement. Forbes & Rigobon (2002) used linear correlation, Bekaert et al. (2009) used factor models and Sander & Kleimeier (2003) used Granger causality. There are, however, only two unique co-movement measures, those proposed by Gerber et al. (2022) and Smyth & Broby (2022). One of the reasons given by scholars to explain excessive co-movement in the literature is an asset class effect (Barberis & Shleifer (2003), DeMarzo et al. (2004) and Basak & Pavlova (2013)). The results from Gerber et al. (2022) and Smyth & Broby (2022), however, show its presence over multiple asset classes.

The literature, as typified by Greenwood & Sosner (2007), tends to refer to excessive co-movement during periods when demand is not being driven by fundamental news. As significant driver of such demand is index funds, who trade on index re-constitutions (Greenwood (2008), Mase (2008) and Dhaene et al. (2012).

There has been growing investigation of co-movement of international markets Shiller (1989), Rua & Nunes (2009), Longin & Solnik (2001), and Brooks & Negro (2006)). The majority of such studies have found that co-movement varies over time. Brooks & Del Negro (2004), for example, showed that co-movement had increased "temporarily" in international markets as a result of the IT bubble in the late 1990's. There are currently three approaches to the measurement of co-movement over time, our method adding a fourth. The first, used by the later study, is to identify this they used a rolling window correlation coefficient. The second approach is to measure the change of co-movement over times using non-overlapping sample periods. The third approach is to use time varying parameter models.

There is also substantial literature on market efficiency. Lim & Brooks (2011) produce a good survey of the evolution of the empirical approaches used by scholars. They find that the vast majority of papers assume that market efficiency is static, an assumption which our results and the aforementioned papers challenge. The speed with which information is accurately reflects in prices is a more recent strand of the literature, Chelley-Steeley (2001), for example, estimate a price adjustment coefficient. This is based on the partial mechanism proposed by Amihud & Mendelson (1987), which incorporates adjustments for noise.

# 4. The Gerber and Smyth-Broby co-movement measures

We now introduce the Gerber family in more detail. The GS is a robust co-movement measure for covariance matrix estimation for the purposes of portfolio construction. It is based on Kendall's Tau but incorporates a noise exclusion zone which eliminates co-movements which fall below a certain threshold in magnitude. The rationale is that such co-movements are sources of noise and will obfuscate any attempt to measure the underlying co-movement connection between statistical variables; in this study between financial return time series. Gerber et al. (2022) demonstrate the efficacy of the Gerber statistic in an analysis which clearly outlines its potential to outperform other commonly used methods of co-movement matrix estimation, historical correlation and shrinkage.

The power of the GS stems from its ability to overcome the main shortcomings of conventional product-moment correlation coefficients. Traditional correlation has no noise filter; all co-movements are taken into account whether they are really sources of noise or sources of co-movement information. This is particularly problematic at both ends of the amplitude scale. Small movements are very plentiful so what they lack in size they make up for in volume. Thus, even though these movements have low impact individually, and are centred on zero, collectively they can contribute substantially to the calculation; sample variation and sample size limitation means that these chance movements do not cancel on aggregate.

Large movements are much less plentiful but much more impactful individually. As with small movements they do not collectively cancel on aggregate. Conventional correlation calculations are subject to the issues of small movements through volume and large movements through size. Pairing such movements across assets means that the resulting co-movement series carries damaging levels of noise.

Problems arising from movements of small amplitude are catered for with the threshold and noise exclusion zone. Problems arising from movements of large amplitude are catered for with the use of a one-size-fits-all count. Thus, a large co-movement, which happens to be noise and not information, is not able to dominate because its impact is limited to a weight of 1. All qualifying co-movements are allocated a weight of 1, hence the term count. As such, the count approach mitigates against noise contributions of large magnitude which can have a distorting effect in conventional correlation calculations.

The Smyth-Broby enhancement, which was first reported in Smyth & Broby (2022), sought to implement design improvements on the original Gerber statistic. The first design amendment relates to the location of the single noise exclusion zone for the Gerber statistic. The second design amendment relates to the use of bespoke contributions rather than counts. Both design considerations are explained in the aforementioned paper.

The GS and SB both assist with the measurement of risk and diversification in a portfolio. They do this by disqualifying those co-movements which are attributable to noise and classifying qualifying co-movements as being in tandem (concordant), inverse (discordant), or exhibiting no discernible relationship (neutral). This allows them to focus on those co-movements that may convey useful information. The GS and SB co-movement matrix elements are structurally identical and are defined analogously as:

$$g_{ij}(t) = \frac{\sum_{\{\mathrm{m}\}} \left(\Delta_{ij}(t_m)_{\mathrm{C}} - \Delta_{ij}(t_m)_{\mathrm{D}}\right)}{\sum_{\{\mathrm{m}\}} \left(\Delta_{ij}(t_m)_{\mathrm{C}} + \Delta_{ij}(t_m)_{\mathrm{D}} + \Delta_{ij}(t_m)_{\mathrm{N}}\right)} \tag{1}$$

where $\{m\}$ is the set of qualifying co-movements in the lookback window $[t-T, t]$. $\{\mathrm{C,D,N}\} \equiv$ {Concordant, Discordant, Neutral} and

$$\Delta_{ij}\left(t_m\right) = \frac{\sqrt{\left(1+|\tilde{r}_i|\right)\left(1+|\tilde{r}_j|\right)}}{1+\left(|\tilde{r}_i|-|\tilde{r}_j|\right)^n} \tag{2}$$

is the real-valued analytic co-movement contribution function, one of two innovations of SB. Setting $\Delta_{ij}\left(t_m\right)=1$ for all $i, j, m$ recovers a counts system akin to GS. Even in the case of counts, GS and SB still differ in concept and implementation due to the redesign of noise exclusion, the other innovation of SB.

As explained, strong-form efficient markets absorb information instantly. If there is an inefficiency in the market then information will be absorbed over time rather than instantly. The phenomenological approximation alluded to above relates to being able to emulate such inefficiency by attaching the notion of value to co-movement data based on age. The central premise is that as information gets absorbed into prices over time, there remains less information to be absorbed as time passes by. Thus, co-movement data from two days ago, for instance, has less predictive power than co-movement data from yesterday. As such it is of lesser value. This can be implemented by applying an information-value penalty to individual co-movements based on age. For simplicity we apply a smooth exponentially decaying weight reduction to all co-movements within a given rolling-window linking its age with perceived predictive value. A detailed account of this is provided in the method section.

Exponential behaviour is characterised by a decay parameter. The decay parameter is de-facto the market information processing speed. Thus, we may choose a range of speed settings and apply them each separately as a speed template. The template is imposed on the market data through the co-movement matrix. A back-test is then be carried out for a given speed setting and outcomes such as the cumulative return, over a period of time, benchmarked against the no-template scenario. The no-template scenario is the default way of performing such analyses. In essence we are tuning over a range of speed (equivalently frequency) settings. If we find a frequency setting which outperforms the no-template scenario then we may conclude that the market is inefficient. We have found the resonance frequency of the market, and the inefficiency is quantified through the associated speed setting.

We shall adopt a convention whereby we refer to an inefficient market as having an information half-life, $t_{1/2}$. This is typically several time-steps.

# 5. Data

We construct a well-diversified portfolio of nine assets, spanning January 1988 to December 2020. This is the same approach utilized in Smyth & Broby (2022). This provides a natural reference point to that publication as well as to the article Gerber et al. (2022) - where the Gerber statistic was introduced originally.

The available parameter space is extensive given the two real-valued threshold parameters $c_1$ (for the confusion zone) and $c_2$ (for the indecision zone). We do not attempt to explore it here as this was done in the aforementioned Gerber family articles. Rather, we confine our analysis to the region $c_1 = 0.5, c_2 = 0.1$. For the data used, the average distance between the centres of the confusion zone $((-\bar{r}_i/s_i, -\bar{r}_j/s_j))$ and the indecision zone $(0, 0)$ is approximately 0.4 standardized units. Thus with settings of $c_1 = 0.5$ and $c_2 = 0.1$, the indecision zone is generally fully encompassed by the confusion zone and thus the confusion zone arbitrates on noise elimination (as is the case in GS). However using a smaller value for $c_2 = 0.1$ has the effect of thinning the neutral branches resulting in a performance advantage which does not have a GS-analogue.

For the sake of clarity, we confine our interest to the counts-based version of SB meaning the penalty parameter $n$ from (2) is redundant for this study. In general, it plays an important role when using the contributions-based version of SB. In this analysis the returns data are monthly and we operate a 24-month lookback rolling window. This equates to a quality parameter of $q \equiv T/N = 24/9$, where $T$ is the number of rows in the rolling window and $N$ is the number of components in the portfolio. The components in the portfolio we consider are listed below,

1. US large capitalization stocks as proxied by the S&P 500 index [SPX]
2. US small capitalization stocks as proxied by the Russell 2000 index [RTY]
3. European, Asia and Far Eastern large and mid capitalization stocks as proxied by the MSCI EAFE index [MXEA]
4. Emerging market large and mid capitalization stocks as proxied by the MSCI Emerging Markets index [MXEF]
5. The US investment grade bond market covering treasuries, government-related and corporate securities, as proxied by the Bloomberg Barclays US Aggregate Bond index [LBUSTRUU]
6. The high yield bond market as proxied by the Bloomberg Barclays US Corporate High Yield Bond index [LF98TRUU]
7. The real estate market as proxied by the FTSE NAREIT all-equity REITS index [FNERTR]
8. Spot Gold [XAU]
9. The commodities market as proxied by the S&P GSCI Goldman Sachs Commodity index [SPGSCI]

We include REIT's, gold and commodities acknowledging that co-movement applies to other asset classes. In this respect, our results corroborate the research of Algieri et al. (2021) who estimated the rolling GS correlation in commodities and showed it to have superior forecasting power.

# 6. Method

By way of method, we further extend the reach of SB as a tool to classify co-movement data by introducing an additional weighting system, designed to penalise older co-movement information as having reduced predictive power on account of its value having already been partially or wholly absorbed by the market. The rationale for this was presented in the previous section and its implementation is demonstrated below via a modification to the contribution function. Specifically, $\Delta_{ij}(t_m)$ is replaced by $\tilde{\Delta}_{ij}(t_m)$, giving

$$\tilde{g}_{ij}(t) = \frac{\sum_{\{m\}} \left(\tilde{\Delta}_{ij}\left(t_m\right)_C - \tilde{\Delta}_{ij}\left(t_m\right)_D\right)}{\sum_{\{m\}} \left(\tilde{\Delta}_{ij}\left(t_m\right)_C + \tilde{\Delta}_{ij}\left(t_m\right)_D + \tilde{\Delta}_{ij}\left(t_m\right)_N\right)} \tag{3}$$

where

$$\tilde{\Delta}_{ij}\left(t_m\right) = \Delta_{ij}\left(t_m\right) e^{-\gamma(m-1)} \tag{4}$$

The information half-life $t_{1/2}$ is related to the exponential decay constant $\gamma$ through $t_{1/2} = \ln(2)/\gamma$. The idea of a half-life is common in exponential treatments of decay phenomena. In this case it represents the time taken for the market to absorb half the value of new information. We find it beneficial to introduce another term, information lifetime $t_\infty$. Information lifetime is the time taken for the market on aggregate to fully absorb new information. Numerically it is the time taken for the exponential weight to have decayed to 0 (to the nearest whole number). Put differently it is the time taken for at least 99.5% of the value of information to have been absorbed by the market. Both terms are useful in conveying market efficiency time frames. Table 1 compares are wide range of values of $\gamma$, $t_{1/2}$ and $t_\infty$.

| decay parameter | Information half-life | Information lifetime |
|---|---|---|
| ($\gamma$) | ($t_{1/2}$) | ($t_\infty$) |
| 0.04 | 18 | 133 |
| 0.5 | 2 | 11 |
| 1 | 1 | 6 |
| 5 | $< 1$ | 2 |
| 10 | $\ll 1$ | $\approx 1$ |
| 50 | $\ll 1$ | $< 1$ |

Table 1: Equivalence of $\gamma$, $t_{1/2}$ and $t_\infty$. $t_{1/2}$ and $t_\infty$ are measured in the same units of periodicity as the returns data.

From Table 1 we can see that a market with an information half-life of two time-units will have an information lifetime of 11 time-units. To be concrete, since we are working in discrete time; for such a market, information arriving during time-unit 1 will have had half of its value extracted by time-unit 3. Straightforward manipulation of (4) will show that a market with an information half-life of five time-units will have an information lifetime of 35 time-units. Working in discrete time and using $m - 1$ in the exponent of (4) means the penalty does not kick-in until time-unit 2. Additionally, using exponential decay rather than formally truncating the number of observations(rows), allows us to preserve the linear algebra structure of the problem. The nominal value of the quality parameter ($q = 24/9$) is preserved throughout.

## 6.1. Portfolio informational efficiency

Central to our contribution is the concept of portfolio informational efficiency (PIE). We define PIE as the time taken for new information affecting the portfolio to be absorbed and reflected in the the prices of its components. The key premise of this publication is that the informational efficiency of a portfolio can be assessed through the co-movement relationship connecting the components. PIE is necessarily an aggregate term since the constituent components of the portfolio will not, in general, have identical informational efficiency. Nor does our approach require them to.

In order to ascertain PIE we impose a fixed age-related $\gamma$-template on all qualifying co-movement data across all asset-pair combinations for each rolling window. We utilize the resulting *SB* co-movement matrix in Markowitz optimization to determine optimum portfolio weights. This process is repeated according to the rebalancing schedule, producing a series of portfolios component weightings spanning the 30-year period over which our study takes place. Finally the portfolio performance is determined out-of-sample in the form of the cumulative return for each of a selection of fitted $\gamma$-templates. This is what is displayed in Figure 1 for each of three investment scenarios (Conservative, Moderate & Aggressive) and for each trial $\gamma$-template used.

INSERT FIGURE 1

The bar labelled $\gamma = 0$ corresponds to having not applied any penalty to the co-movement data and is what is conventionally done in such analyses. The other bars are the result of having imposed a finite propagation speed for information absorption in the market. The other bars therefore represent various levels of informational inefficiency. Alternatively we may think of this as imposing a particular information absorption frequency on the portfolio (or on the market if we are using a market portfolio). The top performing bar will therefore represent the true information absorption frequency (or resonance frequency) of the portfolio; true, that is, given the range of templates applied. If that bar happens to be other than the $\gamma = 0$ bar then we have identified informational inefficiency. The $t_{1/2}$ value of the associated template allows us to quantify the inefficiency.

## 6.2. Atomic informational efficiency

We now present an asset pair approach, which we shall call atomic informational efficiency, as it addresses the combination of assets at the smallest possible level. In this respect, the simplest possible investment portfolio is one consisting of two assets each having a price which is subject to uncertainty at any given time.

Each component in this two-asset portfolio will have a unique informational efficiency. The resulting portfolio will have a unique informational efficiency determined by the informational efficiency of its components. This is the time taken for new information affecting both components to be reflected in their respective prices. Again this is an aggregate term because, as in the case of a multi-asset portfolio, the two assets in this simple portfolio will not have identical informational efficiency. However, a two-asset portfolio constitutes the lowest level of aggregate assumption we can make. As such we may consider the informational efficiency of two-asset portfolios as being the atomic building blocks of informational efficiency.

In the previous section we ascertained PIE for a multi-asset portfolio without disaggregating to the level of asset-pairs. In penalising co-movement data for age, we simply used a single decay rate parameter, $\gamma$, for all co-movement data irrespective of which asset-pair combination we were considering. In effect we were tuning into a resonance frequency for the portfolio as a whole.

In this section we consider an alternative approach in which we first break the multi-asset portfolio down into all possible two-stock portfolios formed from its constituent components. We determine the informational efficiency resonance frequency for each two-asset portfolio separately. This is done by replacing the whole-portfolio decay parameter in (7) with a bespoke decay parameter for each possible two-asset portfolio. Equation (7) becomes

$$\tilde{\Delta}_{ij}\left(t_m\right) = \Delta_{ij}\left(t_m\right) e^{-\gamma_{ij}(m-1)} \tag{5}$$

In effect we tune into each two-asset portfolio separately to determine a resonance frequency $\gamma_{ij}$ for each asset-pair combination. When this has been done the values are stored in a mapping matrix $\Gamma$ whose elements $\Gamma_{ij}$ are $\gamma_{ij}$. We draw from this mapping matrix when building the matrix element $\tilde{g}_{ij}$ of the SB co-movement matrix. Table 2 illustrates the mapping matrix used in this study. The mapping matrix $\Gamma$ should not be confused with the $SB$ co-movement matrix $\tilde{g}$.

| Asset | SPX | RTY | M1EA | EM | XAU | SPGSCI | LF98TRUU | LBUSTRUU | FNERTR |
|---|---|---|---|---|---|---|---|---|---|
| SPX | 0 | 0.03 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| RTY | 0.03 | 0 | 10 | 0 | 0.1 | 10 | 0.03 | 2 | 2 |
| M1EA | 0 | 10 | 0 | 0.8 | 0 | 0.2 | 0.3 | 0 | 10 |
| EM | 0 | 0 | 0.8 | 0 | 10 | 2 | 0.5 | 0.3 | 0 |
| XAU | 0 | 0.1 | 0 | 10 | 0 | 0.5 | 0.2 | 0.8 | 10 |
| SPGSCI | 0 | 10 | 0.2 | 2 | 0.5 | 0 | 10 | 10 | 0.3 |
| LF98TRUU | 0 | 0.03 | 0.3 | 0.5 | 0.2 | 10 | 0 | 10 | 0 |
| LBUSTRUU | 0 | 2 | 0 | 0.3 | 0.8 | 10 | 10 | 0 | 0 |
| FNERTR | 0 | 2 | 10 | 0 | 10 | 0.3 | 0 | 0 | 0 |

Table 2: Symmetric mapping matrix $\Gamma$ for atomic-PIE based on returns data for all two-asset combinations in the nine-asset portfolio. Each element corresponds to the empirically determined resonance frequency $\gamma_{ij}$ for assets $A_i$ and $A_j$.

INSERT FIGURE 2

Figure 2 depicts the impact of adopting this approach rather than the whole-portfolio single-rate approach of the previous section (the $SB_{\gamma=\gamma_P}$ bar). We can see that, for the conservative risk regime, building the portfolio SB co-movement matrix from the atomic level has resulted in enhanced performance (the $SB_{\gamma=\gamma_A}$ bar). We have introduced columns for historical covariance ($HC$) and shrinkage ($SM$) to provide context and allow for performance comparison with $SB$. The atomic approach did not result in an improvement for the moderate or aggressive risk regimes.

In summary, the atomic approach to calculating PIE is to first find a resonance frequency for each possible asset-pair combination. Then build the portfolio SB co-movement matrix from the resulting collection of resonance frequencies and send that co-movement matrix to the Markowitz optimisation. In the more general approach of the previous section we assign a single decay parameter value to the co-movement data for all asset-pair combinations. We subsequently declare the PIE resonance frequency to be the decay parameter value which results in the greatest out-of-sample performance.

# 7. Discussion

We now discuss the practical applications of our approach. We suggest that our technique for identifying and quantifying inefficiency can become part of the optimisation process. Specifically, an initial frequency scan to find the portfolio resonance frequency prior to performing mean-variance optimisation using the associated SB co-movement matrix for the purposes of prediction. The frequency scanning process can also become something that is done habitually to detect change, and re-tune accordingly.

It is clear from our analysis that although a portfolio can be of arbitrary construction, through applying our technique, the efficiency of a market may be tested and the inefficiency quantified by replacing the nominal portfolio with a market portfolio. Indeed, this would be the case with any portfolio whose PIE can be measured.

We make the phenomenological assumption that the informational inefficiency of a portfolio manifests in the co-movement dynamics of the collection of stocks comprising the portfolio. The validity of this assumption is demonstrated empirically and is the key result of the research. We introduce a time-value penalty to co-movement information based on its age and premised on the fact that more recent co-movement data have greater predictive power. Older information has been partially or wholly absorbed and only that component of it yet to be absorbed has predictive value.

We believe our approach also has intuitive appeal. That is, irrespective of the risk scenario used, we should return the same resonance frequency if PIE is a fundamental property of the portfolio. This is evidenced in Figure 1 where performance is optimized for $\gamma = 0.04$ for all three risk regimes, given the collection of $\gamma$-templates used.

As evident from Figure 1, the final two bars on the right for each investment scenario are of equal height. This is evidence of convergence. Convergence in the sense that the informational half-life for the settings involved are both so short (even though one is much shorter than the other), and the concomitant age-penalty so severe than only the first co-movement in the rolling window qualifies for a count. All others are effectively eliminated from a weighting point of view but remain from a numerically point of view preserving the linear algebra structure of the problem. The single effective qualifying co-movement will be categorised as being concordant, discordant or neutral. The denominator of the matrix element fraction in (3) will be 1 and the numerator will be one of {-1,0,1}. This means that the co-movement matrix generated under one of the latter two templates from Figure 1 will be populated with elements taken from {-1,0,1}. It is quite likely, therefore, that statistical incoherence will arise. This is turn will induce numerical incoherence which will manifest as a non-PSD co-movement matrix through the appearance of negative eigenvalues.

In our approach, we have accounted for this possibility by using an alternating projections algorithm to replace a non-PSD co-movement with the nearest PSD co-movement matrix (Higham (2002)). Hence, the relative under-performance in the latter two bars of each scenario is not a result of non-PSD matrices, as this fixed in the code. Rather, it is a direct result of the lack of intelligence associated with utilising just one co-movement value in each rolling window. All results presented in this paper are based on PSD co-movement matrices.

We introduced the concept of atomic informational efficiency whereby a resonance frequency is determined for each and every asset-pair combination comprising the man portfolio. The results of this are passed to a data-driven mapping matrix containing the bespoke collection of resonance frequencies. This matrix is subsequently drawn from for the Markowitz optimisation. We demonstrated in Figure 2 that this approach produces significant improvement for the conservative risk regime. However it did not enhance the outcome for the other risk regimes. This may have been anticipated since this approach is not tapping into a fundamental property of the portfolio. Rather it is considering the portfolio as a data-driven non-linear combination of the atomic informational efficiencies. The data-driven aspect is key and explains why we get different results for differing risk regimes.

The level of enhancement achieved in the conservative risk regime fully warrants the inclusion of atomic-PIE in the quantitative analysis toolkit. Of course, this needs to be factored against the inevitable increase in computational burden associated with atomic-PIE. For small-dimension portfolios such as we have considered in this study, the burden is insignificant, but it scales at approximately $O(n^2)$ with the dimension of the co-movement matrix. Additionally, this approach will result in the increased incidence of non-PSD matrices. Whilst this is easily resolved as the offending matrices are only mildly non-PSD, it does place an additional computational burden.

# 8. Conclusion

In this paper we present a statistic, adapted from the Gerber family of co-movement statistics, designed to measure informational efficiency and place a value on the inefficiency evidenced in equity co-movement. We show how informational efficiency can be quantified in the context of a small portfolio and a designated market proxy based on one of eight asset classes. We further demonstrate how our approach can be used to measure portfolio informational efficiency (PIE), and that the PIE resonance frequency is a fundamental property of the portfolio. We introduced an alternative (high-resolution) approach operating at the fundamental asset-pair level. Atomic-PIE which has the potential to add-value to the more generic calculation of PIE.

Our proposed technique to determine PIE commences its operation at the level of the individual asset-pair. As a result of this, we refer to it as atomic informational efficiency, since this is the fundamental building block for portfolio informational efficiency.

In our approach, we take advantage of the robustness and superior noise elimination properties of the GS and SB measures to establish a co-movement measure of informational efficiency. In doing this, we extend the way SB classifies co-movement data. Our findings can contribute to the long running debate on the EMH by providing a quantitative tool to differentiate between different levels of informational inefficiency in markets.

We add to the literature on both co-variance and EMH. Importantly, our approach is supportive of finance theory and market efficiency whilst reflecting the fact that some risks and market movements are unexplained. We extend what is known about portfolio informational efficiency (PIE) at both the conceptual and the asset pair level.

# References


Algieri, B, Leccadito, A & Toscano, P (2021), 'A time-varying gerber statistic: Application of a novel correlation metric to commodity price co-movements,' *Forecasting*, **3**(2), pp. 339–354.

Amihud, Y & Mendelson, H (1987), 'Trading mechanisms and stock returns: An empirical investigation,' *The Journal of Finance*, **42**(3), pp. 533–553.

Aue, A, Hörmann, S, Horváth, L & Reimherr, M (2009), 'Break detection in the covariance structure of multivariate time series models,' *The Annals of Statistics*, **37**(6B), pp. 4046–4087.

Barberis, N & Shleifer, A (2003), 'Style investing,' *Journal of financial Economics*, **68**(2), pp. 161–199.

Barberis, N, Shleifer, A & Wurgler, J (2005), 'Comovement,' *Journal of financial economics*, **75**(2), pp. 283–317.

Basak, S & Pavlova, A (2013), 'Asset prices and institutional investors,' *American Economic Review*, **103**(5), pp. 1728–58.

Bekaert, G, Hodrick, RJ & Zhang, X (2009), 'International stock return comovements,' *The Journal of Finance*, **64**(6), pp. 2591–2626.

Breymann, W, Dias, A & Embrechts, P (2003), 'Dependence structures for multivariate high-frequency data in finance,' *Quantitative finance*, **3**(1), p. 1.

Brooks, R & Del Negro, M (2004), 'The rise in comovement across national stock markets: market integration or it bubble?' *Journal of Empirical Finance*, **11**(5), pp. 659–680.

Brooks, R & Negro, MD (2006), 'Firm-level evidence on international stock market comovement,' *Review of Finance*, **10**(1), pp. 69–98.

Campbell, JY, Lettau, M, Malkiel, BG & Xu, Y (2001), 'Have individual stocks become more volatile? an empirical exploration of idiosyncratic risk,' *The journal of finance*, **56**(1), pp. 1–43.

Chelley-Steeley, P (2001), 'Opening returns, noise, and overreaction,' *Journal of Financial Research*, **24**(4), pp. 513–521.

Chen, H, Singal, V & Whitelaw, RF (2016), 'Comovement revisited,' *Journal of Financial Economics*, **121**(3), pp. 624–644.

Cheng, PL & Deets, MK (1971), 'Portfolio returns and the random walk theory,' *The Journal of Finance*, **26**(1), pp. 11–30.

Darrat, AF & Zhong, M (2000), 'On testing the random-walk hypothesis: a model-comparison approach,' *Financial Review*, **35**(3), pp. 105–124.

DeMarzo, PM, Kaniel, R & Kremer, I (2004), 'Diversification as a public good: Community effects in portfolio choice,' *The Journal of Finance*, **59**(4), pp. 1677–1716.

Dhaene, J, Linders, D, Schoutens, W & Vyncke, D (2012), 'The herd behavior index: A new measure for the implied degree of co-movement in stock markets,' *Insurance: Mathematics and economics*, **50**(3), pp. 357–370.

Fama, EF (1970), 'Efficient capital markets: A review of theory and empirical work,' *The journal of Finance*, **25**(2), pp. 383–417.

Fama, EF (1991), 'Efficient capital markets: Ii,' *The journal of finance*, **46**(5), pp. 1575–1617.

Forbes, KJ & Rigobon, R (2002), 'No contagion, only interdependence: measuring stock market comovements,' *The journal of Finance*, **57**(5), pp. 2223–2261.

Gerber, S, Markowitz, HM, Ernst, PA, Miao, Y, Javid, B & Sargen, P (2022), 'The gerber statistic: a robust co-movement measure for portfolio optimization,' *The Journal of Portfolio Management*, **48**(3), pp. 87–102.

Greenwood, R (2008), 'Excess comovement of stock returns: Evidence from cross-sectional variation in nikkei 225 weights,' *The review of financial studies*, **21**(3), pp. 1153–1186.

Greenwood, RM & Sosner, N (2007), 'Trading patterns and excess comovement of stock returns,' *Financial Analysts Journal*, **63**(5), pp. 69–81.

Heston, SL, Korajczyk, RA & Sadka, R (2010), 'Intraday patterns in the cross-section of stock returns,' *The Journal of Finance*, **65**(4), pp. 1369–1407.

Higham, N (2002), 'Computing the nearest correlation matrix—a problem from finance,' *Journal of Numerical Analysis*, **22**, pp. 329–343.

Jobson, JD & Korkie, B (1980), 'Estimation for markowitz efficient portfolios,' *Journal of the American Statistical Association*, **75**(371), pp. 544–554.

LeRoy, SF (1989), 'Efficient capital markets and martingales,' *Journal of Economic literature*, **27**(4), pp. 1583–1621.

Levy, M & Roll, R (2010), 'The market portfolio may be mean/variance efficient after all: The market portfolio,' *The Review of Financial Studies*, **23**(6), pp. 2464–2491.

Lim, KP & Brooks, R (2011), 'The evolution of stock market efficiency over time: A survey of the empirical literature,' *Journal of economic surveys*, **25**(1), pp. 69–108.

Longin, F & Solnik, B (2001), 'Extreme correlation of international equity markets,' *The journal of finance*, **56**(2), pp. 649–676.

Malkiel, BG (1989), 'Efficient market hypothesis,' in *Finance*, Springer, pp. 127–134.

Markowitz, HM (1991), 'Foundations of portfolio theory,' *The journal of finance*, **46**(2), pp. 469–477.

Mase, B (2008), 'Comovement in the ftse 100 index,' *Applied Financial Economics Letters*, **4**(1), pp. 9–12.

Parsley, D & Popper, H (2020), 'Return comovement,' *Journal of Banking & Finance*, **112**, p. 105223.

Phoa, W (2013), 'Portfolio concentration and the geometry of co-movement,' *The Journal of Portfolio Management*, **39**(4), pp. 142–151.

Poon, SH, Rockinger, M & Tawn, J (2004), 'Extreme value dependence in financial markets: Diagnostics, models, and financial implications,' *The Review of Financial Studies*, **17**(2), pp. 581–610.

Roll, R (1984), 'A simple implicit measure of the effective bid-ask spread in an efficient market,' *The Journal of finance*, **39**(4), pp. 1127–1139.

Rua, A & Nunes, LC (2009), 'International comovement of stock market returns: A wavelet analysis,' *Journal of Empirical Finance*, **16**(4), pp. 632–639.

Sander, H & Kleimeier, S (2003), 'Contagion and causality: an empirical investigation of four asian crisis episodes,' *Journal of International Financial Markets, Institutions and Money*, **13**(2), pp. 171–186.

Shiller, RJ (1989), 'Comovements in stock prices and comovements in dividends,' *The Journal of Finance*, **44**(3), pp. 719–729.

Smyth, W & Broby, D (2022), 'An enhanced gerber statistic for portfolio optimization,' *Finance Research Letters*.

Veronesi, P (1999), 'Stock market overreactions to bad news in good times: a rational expectations equilibrium model,' *The Review of Financial Studies*, **12**(5), pp. 975–1007.

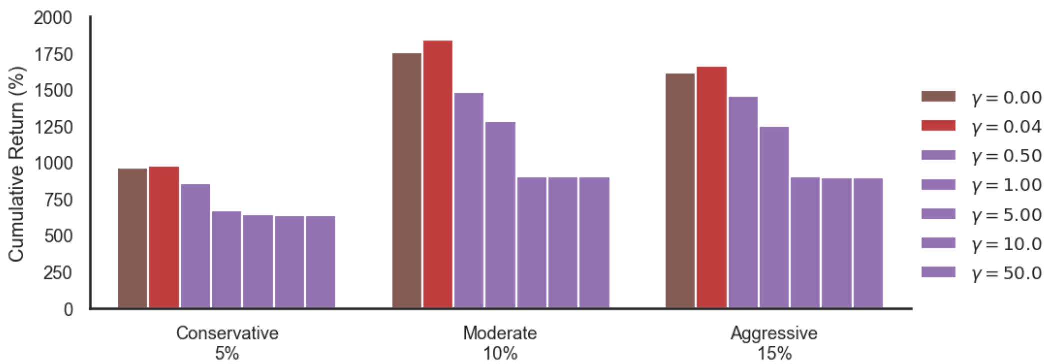


**Figure 1.** Cumulative return portfolio performance for three risk scenarios: Conservative (5%); Moderate (10%); Aggressive (15%); and seven PIE templates ($\gamma = 0, 0.04, 0.5, 1, 5, 10, 50$). The underpinning co-movement statistic is the counts-based Smyth-Broby adjustment to the Gerber statistic with noise exclusion threshold settings $c_1 = 0.5, c_2 = 0.1$.

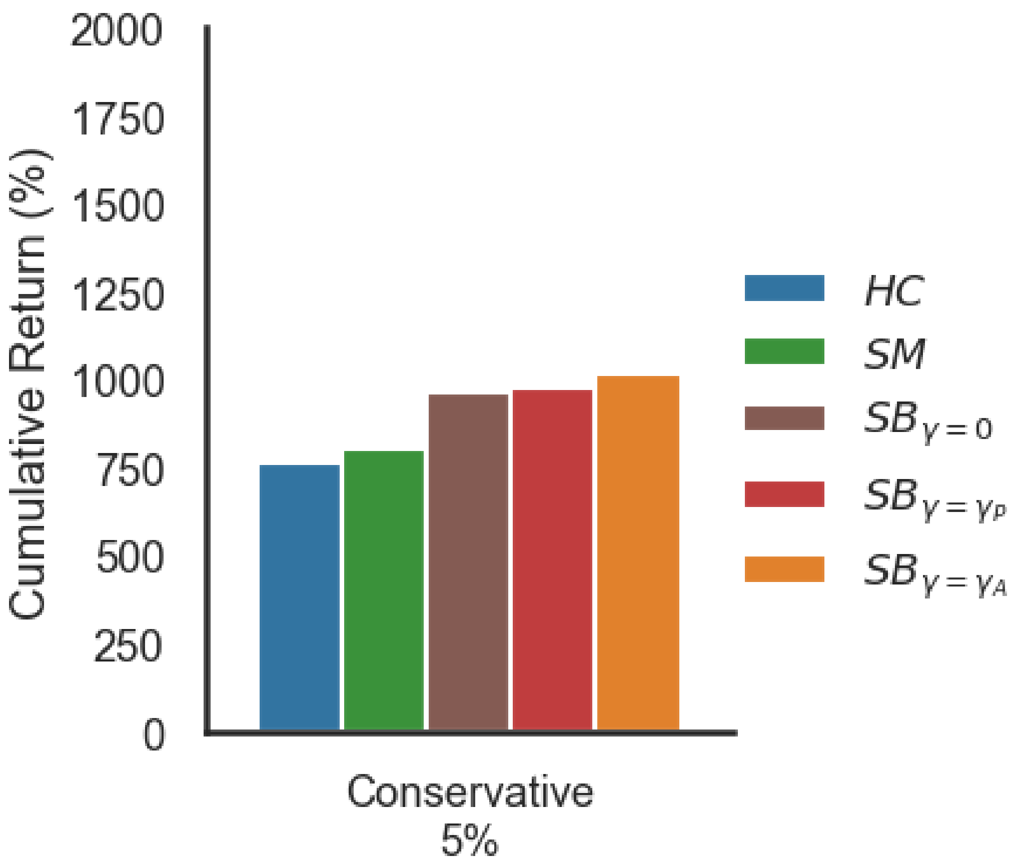


**Figure 2.** Cumulative return portfolio performance for a conservative risk scenario; and five co-movement metrics: ($HC$) historical covariance; ($SM$) shrinkage; ($SB_{\gamma=0}$) without PIE; ($SB_{\gamma=\gamma_P}$) generic PIE; ($SB_{\gamma=\gamma_S}$) atomic-PIE. The $SB$ co-movement statistic used is the counts-based version of the Smyth-Broby adjustment to the Gerber statistic with noise exclusion threshold settings $c_1 = 0.5, c_2 = 0.1$.