PAPER NO. 1 | DIVERSIFICATION & PORTFOLIO RISK | SEPTEMBER 2026

How Much Does Diversification Actually Reduce Portfolio Risk?

Correlation, Covariance, and Volatility Across Six Asset Classes

Rohan Goel | Quanta Student Capital
September 2026

Abstract. This paper asks a simple question with a less-simple answer: does adding more holdings to a portfolio actually reduce its risk, or does what matters is how those holdings move *relative to one another*? We build three portfolios from six ETFs spanning US equities, UK equities, US technology equities, US government bonds, gold, and emerging-market equities (SPY, EWU, QQQ, IEF, GLD, EEM; 1 Jan 2017 – 31 Aug 2026), and decompose each portfolio's volatility exactly via $\sigma_p^2 = \mathbf{w}^\top \Sigma \mathbf{w}$. Doubling the number of equity holdings ($2 \to 4$) cuts volatility by only 1.7 percentage points, while adding a single low-correlation, cross-asset holding ($4 \to 5$) cuts it by a further 4.5 points. We then split the sample into three regimes – the 2020 COVID shock, the 2022 rate-hike shock, and the 2023–25 calm market – and show that equity correlations rise in both crises, but the equity–bond relationship does the opposite of what a single "diversification breaks down in a crisis" story would predict: it goes *more* negative in 2020 (a growth shock) and flips *positive* in 2022 (an inflation shock). The number of assets held is a weak proxy for diversification; the correlation structure – and the type of shock – is what actually governs it.

1 Research Question

Conventional wisdom treats diversification and holding count as roughly interchangeable: more positions, less risk. This paper tests that assumption directly by asking:

How do correlation and asset allocation affect the volatility of a portfolio?

We do not set out to prove that diversification works – we set out to measure *what actually drives it*: is it the number of assets held, or the statistical relationship between them? A concentrated two-asset portfolio, a four-asset "diversified" equity portfolio, and a five-asset cross-asset portfolio are constructed from identical building blocks and compared on that basis.

2 Data & Methodology

2.1 Assets

Six exchange-traded funds were chosen to represent distinct asset classes rather than overlapping bets on the same risk factor:

Ticker	Asset class	Fund
SPY	US equities	SPDR S&P 500 ETF Trust
EWU	UK equities	iShares MSCI United Kingdom ETF
QQQ	US technology equities	Invesco QQQ Trust (Nasdaq-100)
IEF	US government bonds	iShares 7–10 Year Treasury Bond ETF
GLD	Gold	SPDR Gold Shares
EEM	Emerging-market equities	iShares MSCI Emerging Markets ETF

ETFs were used instead of individual stocks so that each series represents a whole asset class rather than company-specific risk. **Sample period:** 1 January 2017 – 31 August 2026 (approx. 9.7 years). **Source:** Portfolio Visualizer's historical return-and-risk engine, computed on monthly total returns (dividends reinvested).[1]

[1] This environment's outbound network access does not reach raw price-data providers (Yahoo Finance, Stooq, etc.) directly, so real historical statistics were sourced through Portfolio Visualizer's own calculation engine rather than re-derived from a locally down-

2.2 Core definitions

For a price series P_t, the log return is

$$r_t = \ln\left(\frac{P_t}{P_{t-1}}\right). \tag{1}$$

Annualised return and volatility scale the periodic mean and standard deviation by the number of periods per year, k (here $k = 12$ for monthly data):

$$R_{\text{annual}} = (1+\bar{r})^k - 1, \qquad \sigma_{\text{annual}} = \sigma_{\text{period}}\sqrt{k}. \tag{2}$$

Risk-adjusted return is measured by the **Sharpe ratio**,

$$S = \frac{R_p - R_f}{\sigma_p}, \tag{3}$$

where R_f is the risk-free rate – here proxied at $R_f \approx 2.0\%$, backed out from the average 3-month T-bill yield over the sample window. **Maximum drawdown** is the worst peak-to-trough decline in cumulative value,

$$\text{MDD} = \min_t \left(\frac{P_t - \max_{s\le t} P_s}{\max_{s\le t} P_s}\right), \tag{4}$$

and **beta** measures an asset's sensitivity to the broad market (m = SPY):

$$\beta_i = \frac{\text{Cov}(R_i, R_m)}{\text{Var}(R_m)} = \rho_{i,m}\frac{\sigma_i}{\sigma_m}. \tag{5}$$

3 Individual Asset Characteristics

Table 1 reports each asset's annualised return, volatility, beta, Sharpe ratio and maximum drawdown over the sample period.

Table 1. Individual asset risk and return, 1 Jan 2017 – 31 Aug 2026

Ticker		Ann. Return	Ann. Vol.	β	Sharpe	Max DD
SPY	US equities	15.38%	15.53%	1.00	0.85	−23.93%
EWU	UK equities	8.89%	15.69%	0.74	0.47	−29.97%
QQQ	US tech	21.28%	19.28%	1.14	0.98	−32.58%
IEF	US govt. bonds	1.07%	6.49%	0.08	−0.18	−23.15%
GLD	Gold	14.58%	15.16%	0.14	0.82	−23.85%
EEM	Emerging markets	9.28%	17.42%	0.77	0.46	−36.52%

Two things stand out before any portfolio is built. First, QQQ's beta of 1.14 confirms it moves *more* than the US market it is nested inside of – it is not really a diversifier away from SPY, it is a leveraged bet on the same factor. Second, IEF and GLD carry the lowest betas (0.08 and 0.14) by a wide margin: on beta alone, these are the two candidates most likely to dampen portfolio-level volatility, which Section 5 confirms directly.

4 Correlation & Covariance Analysis

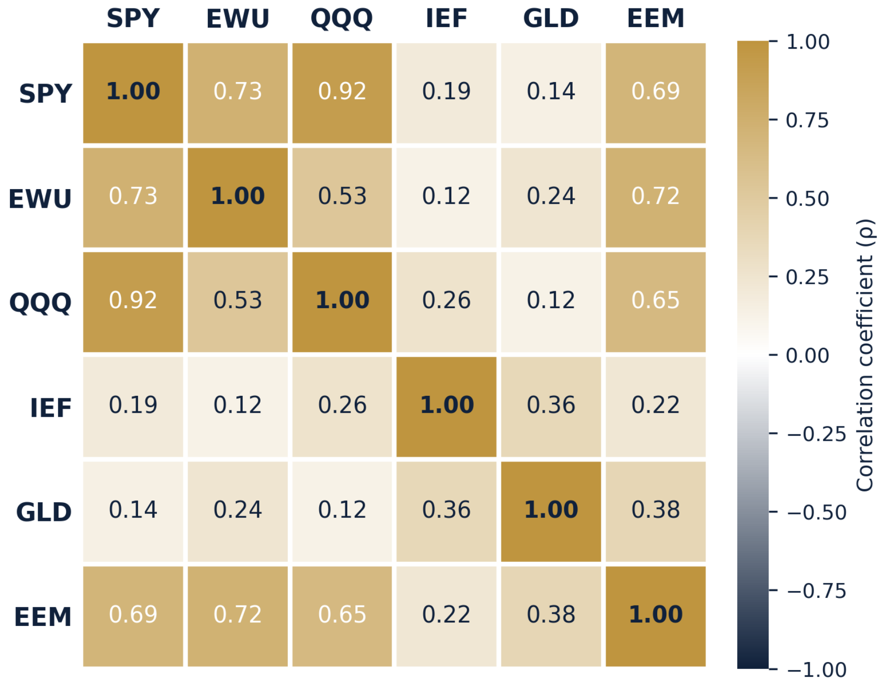
Correlation is the standardised covariance between two return series:

$$\rho_{XY} = \frac{\text{Cov}(X,Y)}{\sigma_X \sigma_Y}, \qquad \text{Cov}(X,Y) = \mathbb{E}\big[(X-\mu_X)(Y-\mu_Y)\big]. \tag{6}$$

Figure 1 shows the full 6×6 correlation matrix of monthly returns over the sample period.

loaded price series. See Section 8 for the methodological implications of this.

Figure 1 — Correlation Matrix of Monthly Returns Jan 2017 – Aug 2026



The matrix is the entire paper in one picture. SPY and QQQ sit at $\rho = 0.92$ – holding both buys almost no reduction in risk relative to holding either alone. SPY and EWU, two *different countries'* equity markets, still sit at $\rho = 0.73$: geography diversifies less than intuition suggests once both are developed-market equities. The genuinely low correlations are IEF against every equity in the set ($\rho \in [0.12, 0.26]$) and GLD against every equity ($\rho \in [0.12, 0.24]$) – these are the pairs doing the actual diversification work.

5 Portfolio Construction

For two assets, portfolio variance expands to

$$\sigma_p^2 = w_1^2\sigma_1^2 + w_2^2\sigma_2^2 + 2w_1w_2\sigma_1\sigma_2\rho_{12}. \tag{7}$$

For n assets this generalises to matrix form, the centrepiece identity of this paper:

$$\boxed{\sigma_p^2 = \mathbf{w}^\top \Sigma\, \mathbf{w}}, \qquad \Sigma_{ij} = \rho_{ij}\,\sigma_i\,\sigma_j, \tag{8}$$

where $\mathbf{w}$ is the $n \times 1$ vector of portfolio weights and Σ is the $n \times n$ annualised covariance matrix built directly from Table 1's volatilities and Figure 1's correlations.

Worked example – Portfolio A

Substituting Portfolio A's weights ($w_{\text{SPY}} = 0.60$, $w_{\text{QQQ}} = 0.40$) and Table 1 / Figure 1's inputs ($\sigma_{\text{SPY}} = 0.1553$, $\sigma_{\text{QQQ}} = 0.1928$, $\rho = 0.92$) into Equation 7:

$$\begin{aligned}
\sigma_p^2 &= (0.60)^2(0.1553)^2 + (0.40)^2(0.1928)^2 + 2(0.60)(0.40)(0.1553)(0.1928)(0.92) \\
&= 0.008683 + 0.005948 + 0.013223 = 0.027853 \\
\sigma_p &= \sqrt{0.027853} = 16.69\%.
\end{aligned}$$

Portfolios B and C are built the same way, just with the 5×5 and 6×6 blocks of Σ in place of the 2×2 block above.

The three portfolios

Table 2. Portfolio construction and results

Portfolio	Weights	Return	Vol. (σ_p)	Sharpe	Avg. ρ
A – Concentrated	60% SPY / 40% QQQ	17.74%	16.69%	0.94	0.92
B – "Diversified" equities	25% each: SPY, EWU, QQQ, EEM	13.71%	14.99%	0.78	0.71
C – Cross-asset	20% each: SPY, EWU, IEF, GLD, EEM	9.84%	10.49%	0.75	0.38

Portfolio return is the weight-averaged realised return of the underlying assets over the identical window; portfolio volatility is the exact matrix result of Equation 7. A useful diagnostic for *how much* diversification a portfolio is actually achieving is the diversification ratio,

$$\mathrm{DR} = \frac{\sum_i w_i \sigma_i}{\sigma_p}, \tag{9}$$

i.e. the weighted-average volatility an investor would bear if every asset moved independently, divided by the volatility they actually bear. $\mathrm{DR} = 1$ means no diversification benefit at all; the further above 1, the more the portfolio's assets are offsetting one another. Portfolio A scores $\mathrm{DR} = 1.02$ – essentially none. Portfolio B, with twice the holdings, only reaches $\mathrm{DR} = 1.13$. Portfolio C, with one more holding than B, reaches $\mathrm{DR} = 1.34$.

Table 3. Holdings vs. realised volatility reduction

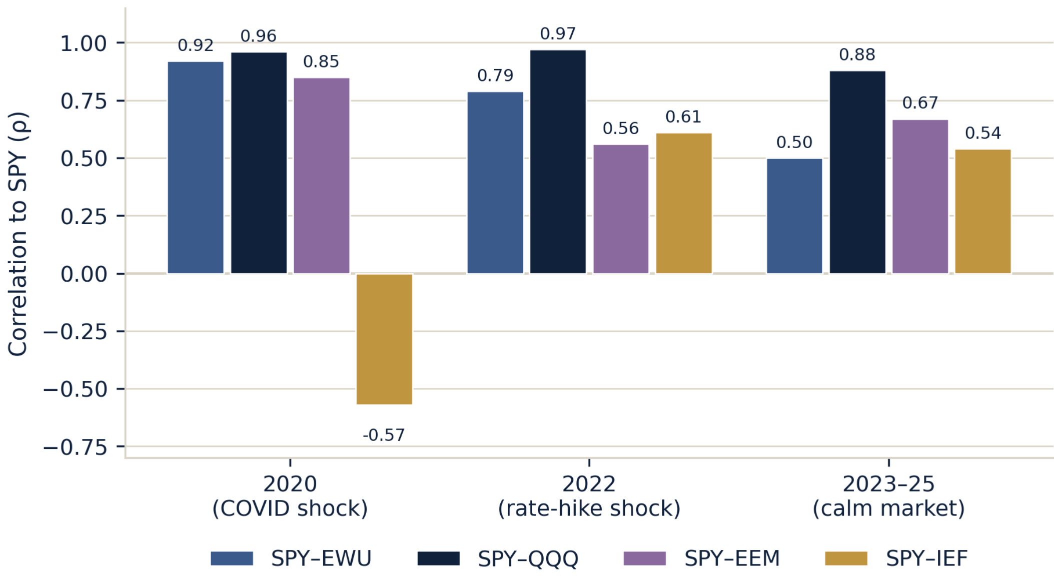
Change	Holdings	$\Delta\sigma_p$	Driver
A $\to$ B	2 $\to$ 4 (equity only)	-1.70 pts	more names, same risk factor
B $\to$ C	4 $\to$ 5 (adds bonds & gold)	-4.50 pts	one low-ρ asset class

Result. Doubling the number of holdings (A$\to$B) barely moves volatility, because every added name shares the same underlying equity risk factor (average pairwise ρ stays above 0.7). Adding a single bond-and-gold exposure (B$\to$C) – one extra holding, not four – produces more than double the volatility reduction, because it lowers the *average correlation* of the book from 0.71 to 0.38. Holding count is a weak predictor of diversification; correlation structure is the real one.

6 Stress-Period Analysis: Do Correlations Change During Crises?

Static, full-period correlations can mask what happens exactly when diversification is needed most. The sample is split into three regimes and the correlation matrix recomputed for each: the 2020 COVID crash, the 2022 rate-hike bear market, and 2023–25 as a calmer comparison window.

Figure 2 — Correlation to SPY Across Market Regimes



Equity-to-equity correlations (SPY–EWU, SPY–QQQ, SPY–EEM) rise in both crisis windows relative to the calm 2023–25 period – the textbook result: in a broad sell-off, geographic and sector diversification within equities compresses toward 1. The more interesting result is SPY–IEF, the equity–bond relationship, which does not behave consistently across the two crises: it falls to -0.57 in 2020 but rises to $+0.61$ in 2022. In 2020, a growth and liquidity shock, government bonds rallied as a flight-to-quality hedge while equities fell – the negative correlation *strengthened* exactly when a hedge was most valuable. In 2022, an inflation and rate-hike shock, both asset classes were repriced by the same rising-discount-rate mechanism, so bonds fell alongside equities – the classic 60/40 hedge failed for a structural reason, not a temporary one.

Result. Diversification benefits are not just weaker in a crisis; *which* assets diversify a portfolio depends on the type of shock. A hedge that worked in the last crisis is not guaranteed to work in the next one if the next one has a different macroeconomic cause.

7 Conclusion & Limitations

Across three portfolios built from the same six building blocks, the number of holdings explained very little of the realised volatility difference; average pairwise correlation explained almost all of it. Doubling equity holdings (Portfolio A $\rightarrow$ B) reduced volatility by 1.7 points; adding one cross-asset holding (B $\rightarrow$ C) reduced it by 4.5 points. The stress-period analysis adds a second layer: even the "good" diversifiers are regime-dependent – government bonds hedged the 2020 shock and failed to hedge the 2022 shock, because the two crises had different macroeconomic drivers.

Limitations. (i) Portfolio-level returns are weight-averaged realised CAGRs rather than a full daily price-path backtest, so compounding/rebalancing drag on the *return* figure (not the volatility figure, which is exact) is only approximated; portfolio-level maximum drawdown was not computed for the same reason. (ii) Correlations are estimated on monthly returns; a daily-return matrix would likely show modestly different short-horizon correlations. (iii) The risk-free rate is a single backed-out approximation ($\approx$ 2.0%) rather than a period-matched daily series. (iv) Historical correlations are not structural constants – as Section 6 shows, they move with the macroeconomic regime, so the exact matrix in Figure 1 is a description of 2017–2026, not a forecast. (v) The choice of start date, ETF proxies, and USD-denominated returns (currency risk is not separated out for EWU or EEM) all affect the precise numbers reported here. (vi) Transaction costs, taxes, and rebalancing costs are excluded throughout. None of this changes the central, structural result – correlation, not headcount, drives diversification – but all of it should temper how literally the specific percentages are read.