What the Stock-Bond Correlation Does Not Measure: Bond Diversification in the Tail, 1962–2026

Philip Kroos*

August 13, 2026

Abstract

The stock-bond correlation turned positive in 2022 and the balanced portfolio lost its diversification. I ask whether the correlation is the right measure of what was lost. Using 16,065 daily observations across 258 quarters of US data, I define the hedging benefit of a bond sleeve as the reduction in portfolio volatility relative to the equity sleeve alone, which depends on the correlation and on the ratio of the two sleeve volatilities and on nothing else. A Shapley decomposition attributes 99 percent of the 2015–19 to 2021–23 deterioration to the correlation channel: equity and bond volatility rose together and left their ratio nearly unchanged. The correlation focus is therefore justified for the average. It is not justified for the tail. Bonds returned +0.03 percent on the worst equity days of the post-2021 quarters against +0.33 percent before, but conditional correlations are mechanically attenuated by conditioning, and against a normality benchmark that removes this bias the post-2021 tail excess is −0.10, the most negative of any regime in the sample. The macro state that the literature uses to explain the unconditional correlation explains it ($R^2 = 0.19$) but not the tail ($R^2 = 0.07$), and in real time, with publication lags imposed, it is beaten decisively by last quarter's realised correlation at a one-quarter horizon (ΔRMSE significant at the one percent level) and only draws level at one year. Correlation-aware allocation improves the Sharpe ratio from 0.68 to 0.79, but the naive rule beats the macro model and the return difference is not distinguishable from zero.

Keywords: stock-bond correlation, diversification, asymmetric dependence, out-of-sample forecasting, asset allocation **JEL:** G11, G12, E44

*philip.kroos@web.de. Replication code and data are at
`github.com/Philip-Kroos/bond-hedge-decomposition`. All errors are my own.

1 Introduction

The balanced portfolio rests on an empirical regularity rather than a theorem. For two decades before 2021, US Treasury returns were negatively correlated with equity returns, so a bond allocation reduced portfolio volatility while contributing return. In 2022 both legs fell together, the correlation turned positive, and the sixty-forty portfolio recorded its worst calendar year in half a century. A large literature has since asked what determines the sign of that correlation, and a persuasive answer exists: it depends on whether inflation or real activity is the dominant source of macroeconomic news [Campbell et al., 2017, 2020].

This paper asks a prior question. The correlation is a summary statistic, and investors hold bonds for a specific purpose, namely to reduce portfolio risk and in particular to cushion equity drawdowns. Whether the correlation is a good measure of how well bonds serve that purpose is not obvious, and it turns out to be true in one respect and false in another.

The empirical difficulty is that the natural object, the co-movement of bonds and equities conditional on equities falling, cannot be estimated by computing a correlation on the bad days. Selecting a subsample by the realisation of one variable truncates its variance and attenuates the measured correlation toward zero even when the underlying dependence is perfectly symmetric. The point is familiar in the contagion literature from Boyer et al. [1997] and Forbes and Rigobon [2002], and it is quantitatively large here: a quarter whose true correlation is -0.40 delivers an expected correlation of only -0.20 in its lowest equity quartile under joint normality. Roughly half the apparent weakening of the hedge in drawdowns is an artefact of the conditioning.

I address this in three steps. First, I define the object of interest so that it isolates what a portfolio actually experiences. For a portfolio with weight w in equities, compare its volatility to that of the equity sleeve alone. The ratio depends on the correlation ρ and on k, the bond sleeve's volatility relative to the equity sleeve's, and on nothing else; I call $H = 1 - \sigma_p/(w\sigma_e)$ the hedging benefit. Second, I benchmark every conditional correlation against the value a bivariate normal with the same unconditional correlation would produce under the same conditioning rule, obtained by simulation because the empirical rule conditions on a within-sample quantile with finitely many observations. Third, I evaluate the macro predictors of the literature under a real-time rule, with an expanding estimation window and an explicit publication lag, against the benchmarks a practitioner would actually use.

The data are 16,065 daily observations from January 1962 to June 2026. Equity excess returns are the Fama-French market factor; bond returns are constructed from the ten-year constant-maturity yield by the par-bond approximation with duration and convexity evaluated at the previous day's yield. The construction reproduces known outcomes closely, delivering a 2022 total return of -16.4 percent against roughly -16 percent for the corresponding index.

Four results follow. *First*, the correlation focus is right for the average. A Shapley decomposition of the change in H between 2015–19 and 2021–23 assigns -0.148 of the total -0.149

to the correlation channel. Bond volatility rose by roughly half, but so did equity volatility, and the ratio that enters H barely moved. The widespread claim that rising bond volatility drove the diversification loss is not supported.

Second, the correlation focus is wrong for the tail. On the five worst equity days of each quarter, bonds returned +0.33 percent on average during 1998–2020 and +0.03 percent since 2021: much reduced, but not negative, and far from the −0.35 percent of the 1982–97 disinflation. Against the normality benchmark, the post-2021 tail excess is −0.10, the most negative figure in the sample, and the pooled downside excess is −0.035 with a bootstrap p value of 0.009. Conditional on equities falling, bonds co-moved *less* than symmetric dependence implies, and more so after 2021 than before. The unconditional correlation overstates the damage to the hedge.

Third, the macro mechanism reaches the average but not the tail. The standard predictors explain 19 percent of the variation in the unconditional correlation, 12 percent of the downside correlation, 7 percent of the tail correlation, and essentially none of the tail excess ($R^2 = 0.001$). Whatever generates the asymmetry is not the inflation-growth mix.

Fourth, the in-sample fits do not survive real time in the way a reader of that literature might expect. The macro model beats the historical mean at every horizon, with Clark-West statistics above four. Against last quarter's realised correlation it loses decisively at one quarter, with a Diebold-Mariano statistic of 4.95, and only draws level at four quarters. In a portfolio, a correlation-aware rule raises the Sharpe ratio from 0.68 to 0.79 and cuts the maximum drawdown from 27 to 20 percent, but the rule that uses last quarter's correlation beats the one that uses the macro forecast, and the return difference against the static portfolio is not distinguishable from zero (bootstrap $p = 0.62$).

Contribution. The paper contributes to three strands and claims none of the mechanism.

To the literature on the determinants of the stock-bond correlation [Baele et al., 2010, Campbell et al., 2017, 2020], it adds a measurement point and a limit. The measurement point is that the portfolio-relevant object is H, not ρ, and that the two can move apart; here they do not, which is itself worth establishing rather than assuming. The limit is that the mechanism explains the unconditional moment and not the conditional one.

To the literature on asymmetric dependence, it applies an existing correction to a question where it has not been applied and where it reverses the sign of the answer. The benchmark used here is the same idea as the H statistic of Ang and Chen [2002], who compare observed conditional correlations with those a symmetric distribution would produce, applied to equity portfolios rather than across asset classes; Longin and Solnik [2001] pursue the same question with extreme value theory.

To the literature on out-of-sample return and moment predictability [Welch and Goyal, 2008, Campbell and Thompson, 2008], it adds the stock-bond correlation as an object, with the finding that a persistent target makes the random walk a demanding benchmark that the published in-sample fits do not have to clear.

The paper is not causal and does not present itself as such. There is no natural experiment for macroeconomic volatility regimes, and a design claiming one would not be credible. The claim is conditional prediction under a stated information set.

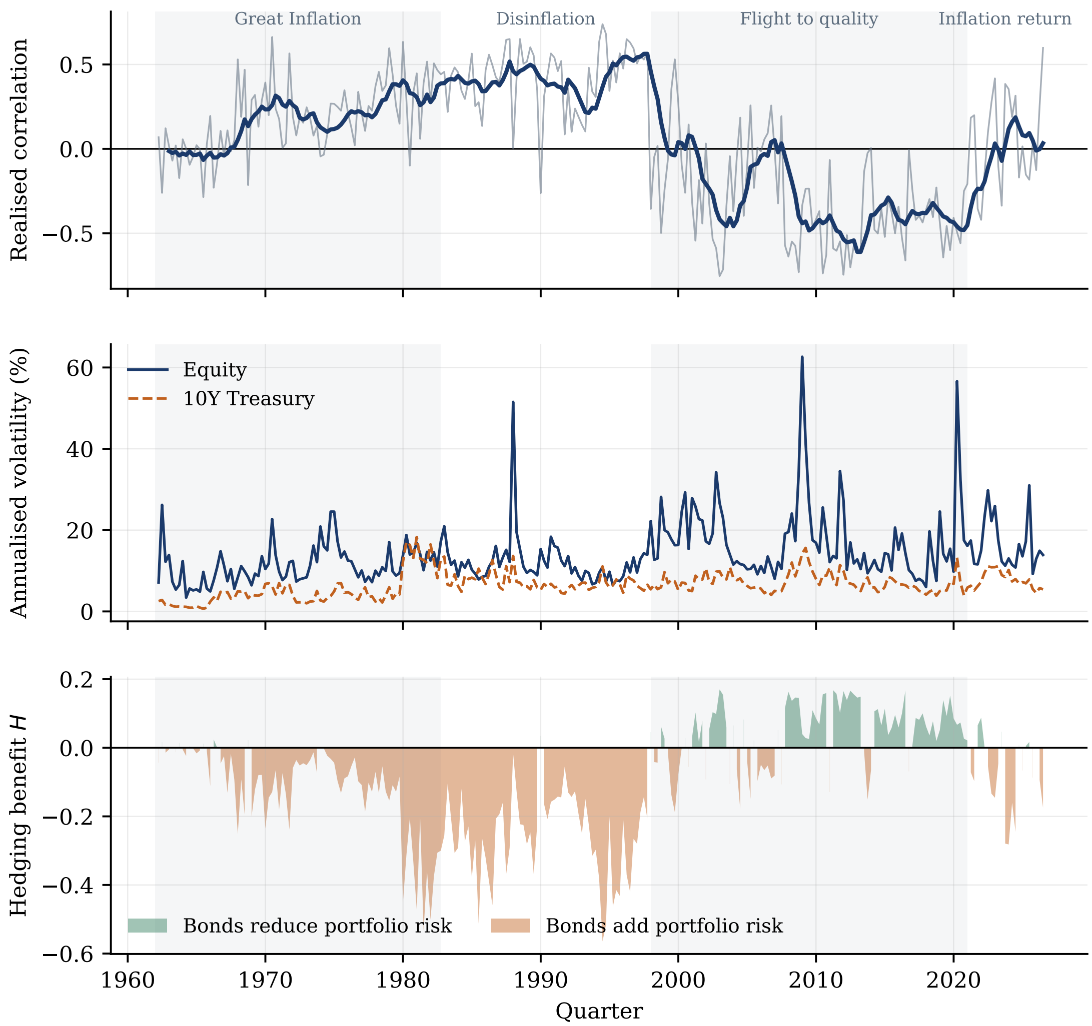

Figure 1: **Sixty years of the hedge.** Realised quarterly correlation, annualised volatilities, and the hedging benefit H of equation (3). $H > 0$ means the bond sleeve lowers portfolio volatility. It is negative in 89 percent of Great Inflation quarters and 97 percent of disinflation quarters: the two decades on which the balanced portfolio's reputation rests are the exception in this sample, not the rule.

Section 2 sets out the framework, Section 3 the data, Section 4 the empirical strategy. Section 5 reports results, Section 6 robustness, Section 7 portfolio implications, and Section 8 concludes.

2 Framework

2.1 What a bond sleeve does to portfolio risk

Let σ_e and σ_b denote equity and bond return volatility and ρ their correlation. A portfolio with weight w in equities has variance

$$\sigma_p^2 = w^2\sigma_e^2 + (1-w)^2\sigma_b^2 + 2w(1-w)\rho\sigma_e\sigma_b. \tag{1}$$

The relevant comparison is not to some undiversified alternative but to what the investor would hold without the bond sleeve, namely the equity sleeve on its own. Dividing by $w\sigma_e$ and writing

$$k \;=\; \frac{1-w}{w}\cdot\frac{\sigma_b}{\sigma_e} \tag{2}$$

for the bond sleeve's volatility relative to the equity sleeve's gives

$$\left(\frac{\sigma_p}{w\sigma_e}\right)^2 = 1 + k^2 + 2k\rho, \qquad H \;\equiv\; 1 - \frac{\sigma_p}{w\sigma_e}. \tag{3}$$

Equation (3) is an identity, not a model, and it has two useful implications. The hedging benefit depends on exactly two quantities, ρ and k, so any statement about diversification that conditions only on the correlation is incomplete. And $H > 0$ requires $k < -2\rho$, so a bond sleeve lowers total portfolio volatility only when the correlation is negative and bonds are not too volatile relative to equities. Under a sixty-forty split with $\sigma_b/\sigma_e = 0.45$, which is close to the post-1998 average, $H > 0$ requires $\rho < -0.15$.

2.2 Decomposing a change in the hedge

H is a function of three primitives. To attribute a change between two periods to each, I use the Shapley value: move each primitive from its initial to its final value, average the marginal contribution over all 3! orderings. The decomposition is exact and order independent by construction, and the residual reported in Table 2 is a numerical check.

2.3 Why a tail correlation cannot be read directly

Let (X, Y) be jointly normal with correlation ρ and consider $\rho^{\text{tail}} = \text{corr}(X, Y \mid X \leq q_\tau)$. Conditioning truncates the variance of X; under normality the conditional correlation is strictly smaller in magnitude than ρ, with the attenuation increasing as τ falls. Comparing an estimated ρ^{tail} to ρ therefore tests nothing about dependence.

The correct null is the value implied by symmetric dependence at the same unconditional correlation and under the same conditioning rule. I compute it by simulation rather than the closed form, because the empirical rule conditions on a within-sample quantile estimated from about sixty observations, and the closed form assumes a fixed population threshold. Define

$$\text{tail excess}_t \;=\; \hat{\rho}_t^{\text{tail}} - \mathbb{E}\big[\hat{\rho}^{\text{tail}} \mid \rho = \hat{\rho}_t,\ n = n_t\big]\,, \tag{4}$$

with the expectation taken over 4,000 simulated bivariate normal samples of the same length and under the same conditioning rule. The last clause is not decorative. The downside statistic conditions on a negative equity return, not on the lower half, and the realised share of negative days ranges from 0.27 to 0.71 across quarters; benchmarking it at a fixed one half would compare each quarter with a simulation that truncated at a different point than the data did. A positive tail excess means bonds co-move with equities more in drawdowns than symmetric dependence implies, which is what the phrase "the hedge fails when you need it" asserts. Figure 2(a) shows how large the correction is.

3 Data

The sample runs from 3 January 1962, the first daily constant-maturity Treasury quote, to 30 June 2026, and comprises 16,065 trading days in 258 complete quarters.

Equities. Daily market excess returns from the Fama-French research factors, which are CRSP value-weighted returns net of the one-month Treasury bill.

Bonds. Daily total return indices for constant-maturity Treasuries are not freely available back to 1962, so the return is constructed from the yield,

$$r_t^b \;=\; \frac{y_{t-1}}{252} - D(y_{t-1})\,\Delta y_t + \tfrac{1}{2}C(y_{t-1})\,\Delta y_t^2, \tag{5}$$

where D and C are the modified duration and convexity of a ten-year par bond, evaluated at the previous day's yield and computed by central differences on the exact price function. Convexity is retained because it is not negligible on large yield days and 2022 had many. Both legs are converted to excess returns over the same bill rate; leaving the cash rate in both would induce mechanical positive correlation in the 1979–82 period, when it moved by hundreds of basis points.

The construction is validated against known outcomes. It delivers a 2022 total return of −16.4 percent, close to the roughly −16 percent of the ten-year index, and it ranks 2022, 2009, 2013, 1999 and 1994 as the five worst years, which is the ranking practitioners would name. Appendix Table A1 reports the full comparison.

Macro state. FRED-MD [McCracken and Ng, 2016]. Following the mechanism in Campbell et al. [2020], the predictors are the ratio of rolling inflation volatility to rolling industrial production growth volatility, their rolling correlation, and the level of inflation, all over sixty-month windows. Each series is stamped to a quarter only after a two-month publication lag, so a forecast formed at the end of quarter t uses only months whose releases had occurred. Revisions are not undone; see Section 6.

Quarterly moments. Realised volatilities and correlations are computed from daily observations within each quarter, about 63 per quarter. This matters for the forecasting exercise: unlike rolling windows over monthly returns, the target does not overlap across observations, so predictive regressions are not evaluated against a mechanically persistent dependent variable.

4 Empirical strategy

Descriptive regimes. Table 1 splits the sample into four periods chosen ex ante from the monetary history and not from the data: the Great Inflation to 1982, the disinflation to 1997, the flight-to-quality era to 2020, and the return of inflation from 2021. Nothing in the estimation conditions on these labels.

Inference. The realised correlation is persistent and the effective sample is much smaller than 258. Standard errors on regressions use Newey-West with four lags; tests on means of persistent series use a moving-block bootstrap with eight-quarter blocks and 4,999 replications.

Forecast evaluation. Expanding window from 1990, so that the first origin has 28 years of training. Three benchmarks: the historical mean, an AR(1), and the random walk, meaning last quarter's realised value. I report the Campbell and Thompson [2008] out-of-sample R^2 against each, Diebold and Mariano [1995] statistics with the Harvey et al. [1997] small-sample correction against the random walk, and Clark and West [2007] statistics against the historical mean, since the macro model nests it and Diebold-Mariano is known to under-reject in that case.

What is not claimed. The coefficient on a macro predictor is a conditional expectation, not an effect. No causal interpretation is offered, and no trading rule is presented with a Sharpe ratio detached from the number of specifications examined.

5 Results

5.1 Sixty years of the hedge

Figure 1 shows the three quantities. The correlation is positive through the 1960s, strongly so through the disinflation, sharply negative from 1998, and back to approximately zero after 2021.

Volatility of both legs rises after 2021. The hedging benefit H is negative in 89 percent of Great Inflation quarters, 97 percent of disinflation quarters, 28 percent of flight-to-quality quarters and 64 percent of quarters since 2021. The two decades on which the balanced portfolio's reputation rests are the exception in this sample, not the rule.

5.2 What actually changed in 2021–23

Table 2 and Figure 3 report the decomposition. Between 2015–19 and 2021–23, H fell from $+0.072$ to -0.076. Of that -0.149, the correlation channel accounts for -0.148, bond volatility for -0.018, and equity volatility for $+0.017$. Bond volatility did rise, from 6.0 to 8.8 percent annualised, but equity volatility rose from 12.9 to 17.6 percent, and it is the ratio that enters (3).

This rejects the hypothesis I fixed before estimating, that the volatility channel would dominate. It also disciplines a common claim: the loss of diversification was a correlation event, not a bond-volatility event.

5.3 The tail

Table 3 reports the conditional statistics. Read naively, they support the standard story only weakly: the tail correlation averages 0.050 below the unconditional one, and the downside correlation 0.057 below.

The normality benchmark changes the reading. Figure 2(a) plots the simulated benchmark against the unconditional correlation and shows that roughly half the attenuation is mechanical. Netting it out, the pooled tail excess is -0.027 with a bootstrap interval of $[-0.053, +0.003]$, and the downside excess is -0.035 with interval $[-0.057, -0.013]$ and $p = 0.009$. Conditional on equities falling, bonds co-move less than symmetric dependence implies. The direction is the opposite of what the phrase "the hedge fails in the tail" asserts.

The regime pattern is sharper still. The tail excess is $+0.041$ in the disinflation, -0.022 in the Great Inflation, -0.060 in the flight-to-quality era, and -0.101 since 2021, the most negative in the sample. Figure 5 makes the same point without a correction: on the five worst equity days of each quarter, bonds returned -0.35 percent on average in the disinflation, $+0.33$ percent in the flight-to-quality era, and $+0.03$ percent since 2021. The crash hedge weakened after 2021; it did not invert, and it remains far better than in any previous high-inflation period.

5.4 How far the macro mechanism reaches

Table 4 reports in-sample fits. The macro state explains 19 percent of the unconditional correlation, with the inflation-growth volatility ratio and the inflation level significant and their correlation not. It explains 12 percent of the downside correlation, 7 percent of the tail correlation, and 0.1 percent of the tail excess.

The gradient is the result. The mechanism is a theory of the average co-movement. The asymmetry that determines whether bonds cushion a drawdown is orthogonal to it, at least as measured here, and its source is left open.

5.5 Real time

Table 5 and Figure 4 report the forecast evaluation. Against the historical mean the macro model wins clearly and at every horizon: out-of-sample R^2 between 0.20 and 0.24 for the unconditional correlation, with Clark-West statistics above six.

Against last quarter's realised correlation the picture reverses. At one quarter the out-of-sample R^2 is -1.29 and the Diebold-Mariano statistic 4.95, so the random walk wins decisively. At two quarters the gap narrows and at four quarters it closes ($p = 0.41$). For the tail correlation the macro model and the random walk are indistinguishable at every horizon.

The reading is not that the macro state is uninformative. It is that the target is persistent enough that the sophisticated forecast has to beat a very cheap one, and at the horizon on which allocation committees usually meet, it does not. The published in-sample fits do not face this comparison.

6 Robustness

Table 7 summarises. The decomposition result is not a weighting artefact: at $w \in \{0.5, 0.6, 0.7\}$ the post-2021 H is -0.140, -0.069 and -0.033, all negative, and the correlation share of the change stays above 99 percent. Moving the tail threshold from the lowest quartile to the lowest third leaves the conclusion intact.

At the lowest decile the estimate becomes unusable, averaging $+0.22$, because a correlation computed on six observations is dominated by estimation noise and the normality benchmark cannot rescue an estimator with that little information. I report it rather than omitting it, and I do not use it.

Three limitations are not repaired by robustness checks. Data revisions are not undone: the publication lag is imposed but a genuine vintage exercise would need ALFRED, and the direction of the resulting bias is toward overstating real-time performance. The bond return is an approximation, though the 2022 validation suggests the error is small relative to the effects discussed. And the four descriptive regimes contain effectively three sign changes, so the number of independent regime events is small however many quarters are stacked.

7 Portfolio implications

Table 6 reports a mean-variance allocation with fixed expected returns, so the exercise isolates the value of knowing the second moment. Between 1990 and 2025 the static sixty-forty portfolio returns 6.40 percent at 9.43 percent volatility, a Sharpe ratio of 0.68 and a maximum drawdown of 26.8 percent. Using last quarter's realised correlation gives 5.73 percent at 7.28 percent, a Sharpe ratio of 0.79 and a drawdown of 19.8 percent. The macro forecast gives 0.75 and 21.3 percent.

Two things follow. Correlation awareness helps, and it helps through risk rather than return; the dynamic rules earn less and are penalised less. And the ranking matches the forecast evaluation, with the cheap rule ahead of the expensive one.

The gain should not be oversold. The return difference against the static portfolio is -0.67 percentage points a year with a bootstrap p value of 0.62, so a reader who wants the improvement demonstrated at conventional significance will not find it here. Turnover averages 13 percentage points of equity weight per quarter, which is material for a taxable investor and is charged at five basis points one way in the figures above.

8 Discussion and conclusion

Three statements survive the exercise. The loss of diversification after 2021 was a correlation event, not a bond-volatility event: of the -0.149 change in the hedging benefit, the correlation channel accounts for -0.148. This is contrary to a hypothesis I held before estimating and to a common claim. The correlation nonetheless overstates the damage to the hedge, because conditional on equity drawdowns bonds co-moved less than symmetric dependence implies, and more so after 2021 than in any earlier period. And the macro mechanism that explains the average correlation neither explains the tail nor survives a real-time comparison against last quarter's realised value at horizons shorter than a year.

The practical implication is narrower than the literature's framing suggests. An allocator who wants a view on next quarter's correlation should use last quarter's, and an allocator who wants a view on the coming year can use the macro state without loss. Neither should infer from a positive unconditional correlation that bonds have stopped cushioning drawdowns, because in this sample they have not.

Two questions follow. The asymmetry is real, is largest exactly when the unconditional correlation is least favourable, and is unexplained by the macro state; identifying what generates it is the natural next step, and the flight-to-quality and monetary-response channels make different predictions about its behaviour around policy events. And the result that a persistent target defeats a structural forecast at short horizons is unlikely to be specific to this moment; whether it holds for other second moments used in allocation is testable with the same design.

References

Andrew Ang and Joseph Chen. Asymmetric correlations of equity portfolios. *Journal of Financial Economics*, 63(3):443–494, 2002.

Lieven Baele, Geert Bekaert, and Koen Inghelbrecht. The determinants of stock and bond return comovements. *Review of Financial Studies*, 23(6):2374–2428, 2010.

Brian H. Boyer, Michael S. Gibson, and Mico Loretan. Pitfalls in tests for changes in correlations. International Finance Discussion Paper 597, Board of Governors of the Federal Reserve System, 1997. Revised March 1999.

John Y. Campbell and Samuel B. Thompson. Predicting excess stock returns out of sample: Can anything beat the historical average? *Review of Financial Studies*, 21(4):1509–1531, 2008.

John Y. Campbell, Adi Sunderam, and Luis M. Viceira. Inflation bets or deflation hedges? the changing risks of nominal bonds. *Critical Finance Review*, 6(2):263–301, 2017.

John Y. Campbell, Carolin Pflueger, and Luis M. Viceira. Macroeconomic drivers of bond and equity risks. *Journal of Political Economy*, 128(8):3148–3185, 2020.

Todd E. Clark and Kenneth D. West. Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics*, 138(1):291–311, 2007.

Francis X. Diebold and Roberto S. Mariano. Comparing predictive accuracy. *Journal of Business and Economic Statistics*, 13(3):253–263, 1995.

Kristin J. Forbes and Roberto Rigobon. No contagion, only interdependence: Measuring stock market comovements. *Journal of Finance*, 57(5):2223–2261, 2002.

David Harvey, Stephen Leybourne, and Paul Newbold. Testing the equality of prediction mean squared errors. *International Journal of Forecasting*, 13(2):281–291, 1997.

François Longin and Bruno Solnik. Extreme correlation of international equity markets. *Journal of Finance*, 56(2):649–676, 2001.

Michael W. McCracken and Serena Ng. Fred-md: A monthly database for macroeconomic research. *Journal of Business and Economic Statistics*, 34(4):574–589, 2016.

Ivo Welch and Amit Goyal. A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies*, 21(4):1455–1508, 2008.

Tables and Figures

Table 1: Descriptives. Bonds returned +0.33% on the worst equity days in 1998–2020 and +0.03% since 2021 — weaker, but far from the −0.35% of the disinflation

	Great Inflation 1962–82	Dis-inflation 1982–97	Flight to quality 1998–20	Inflation return 2021–26
Equity volatility (%)	11.323	12.229	17.465	16.473
Bond volatility (%)	5.028	7.153	7.238	7.795
Correlation ρ	0.182	0.410	-0.312	0.060
ρ on down days	0.101	0.264	-0.277	-0.042
ρ in lowest equity quartile	0.073	0.262	-0.226	-0.068
benchmark under normality	0.096	0.221	-0.166	0.033
tail excess	-0.022	0.041	-0.060	-0.101
Volatility ratio k	0.294	0.431	0.312	0.337
Hedging benefit H	-0.114	-0.248	0.042	-0.077
Bond return, 5 worst days (%)	-0.152	-0.347	0.332	0.030
Share of quarters with $H < 0$	0.892	0.967	0.283	0.636
Quarters	83	61	92	22

Quarterly means. Volatilities annualised. k is the bond sleeve's volatility relative to the equity sleeve's at $w = 0.6$; H the hedging benefit of equation (3). "Bond on worst days" is the average daily bond excess return on the five worst equity days of each quarter, in decimals.

Table 2: Decomposition of the change in H. The correlation channel accounts for −0.148 of the total −0.149: this was not a bond-volatility event

	contribution
σ_e	0.016
σ_b	-0.018
ρ	-0.148
total	-0.149
residual	0.000

Table 3: Conditional co-movement. After the normality correction the excess is *negative*: bonds co-move less in drawdowns than symmetric dependence implies

	Mean	5%	95%	p	N
$\rho^{\text{tail}} - \rho$	-0.050	-0.090	-0.005	0.063	258
$\rho^{\text{down}} - \rho$	-0.057	-0.091	-0.023	0.007	258
Tail excess over benchmark	-0.027	-0.053	0.003	0.144	258
Downside excess over benchmark	-0.035	-0.057	-0.013	0.009	258

Moving-block bootstrap, eight-quarter blocks, 4,999 replications. The first two rows compare conditional to unconditional correlations directly and are reported only to show what the uncorrected comparison would suggest.

Table 4: In-sample fit. The macro state explains 19 percent of the unconditional correlation and 0.1 percent of the tail excess

	R^2	N	Inflation/growth vol ratio	Inflation-growth corr.	Inflation level
ρ	0.186	256	-0.641 (0.021)	-0.007 (0.969)	+0.053 (0.000)
ρ^{down}	0.124	256	-0.563 (0.006)	-0.032 (0.810)	+0.035 (0.000)
ρ^{tail}	0.070	256	-0.370 (0.073)	-0.023 (0.852)	+0.030 (0.001)
Tail excess	0.001	256	-0.019 (0.874)	-0.021 (0.775)	+0.002 (0.665)
H	0.145	256	+0.104 (0.323)	-0.056 (0.417)	-0.023 (0.000)

Newey-West standard errors, four lags; p values in parentheses.

Table 5: Out-of-sample evaluation. The macro model beats the historical mean everywhere and loses to last quarter's realised value at one quarter

	R^2_{oos} vs mean	R^2_{oos} vs RW	DM	p(DM)	CW	p(CW)	N
ρ, $h = 1$	0.202	-1.292	4.947	0.000	6.199	0.000	143
ρ, $h = 2$	0.227	-0.577	2.242	0.027	6.543	0.000	142
ρ, $h = 4$	0.243	-0.192	0.820	0.414	6.724	0.000	140
ρ^{tail}, $h = 1$	0.099	0.003	-0.022	0.982	4.493	0.000	143
ρ^{tail}, $h = 2$	0.103	0.204	-1.358	0.177	4.553	0.000	142
ρ^{tail}, $h = 4$	0.097	0.069	-0.460	0.646	4.322	0.000	140

Expanding window from 1990 with a two-month publication lag on macro data. Positive Diebold-Mariano favours the random walk; positive Clark-West favours the macro model over the historical mean.

Table 6: Allocation. Correlation awareness raises the Sharpe ratio from 0.68 to 0.79 and cuts drawdown from 27 to 20 percent — but the cheap rule beats the expensive one

	Mean	Volatility	Sharpe	Max drawdown	N
Static 60/40	0.064	0.094	0.679	-0.268	142
Correlation-aware, last quarter	0.057	0.073	0.786	-0.198	142
Correlation-aware, macro model	0.055	0.074	0.748	-0.213	142

Table 7: Robustness. The correlation share of the decomposition stays above 99 percent at every weight; the decile threshold is reported and not used

	value
H post, w=0.50	-0.140
H post, w=0.60	-0.069
H post, w=0.70	-0.033
mean tail corr, q=0.10	0.219
mean tail corr, q=0.20	-0.010
mean tail corr, q=0.33	-0.005
Shapley rho share	0.992

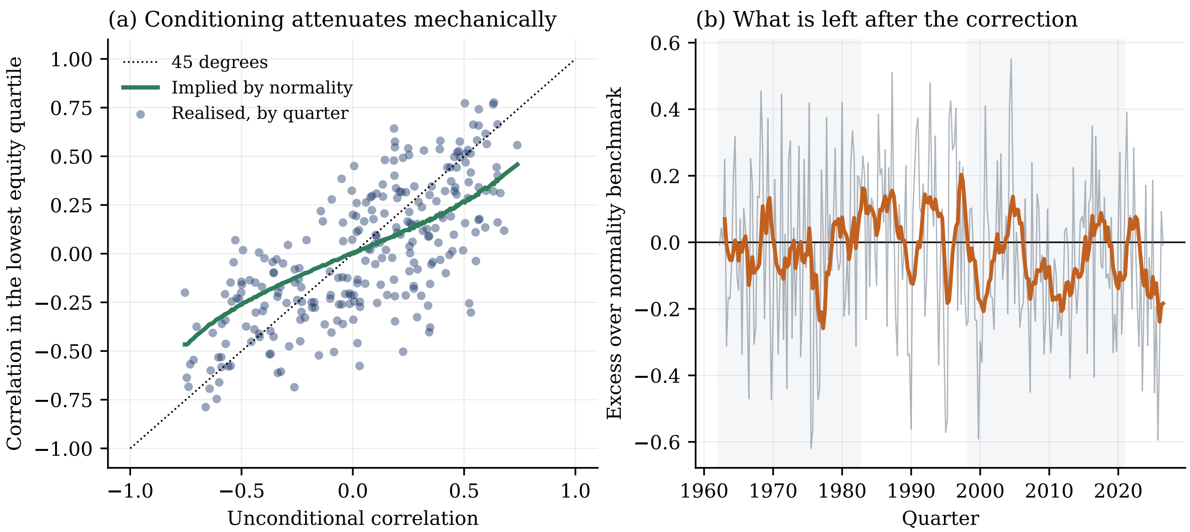

Figure 2: Panel (a): the correlation in the lowest equity quartile against the unconditional correlation, with the value implied by joint normality. Roughly half the apparent attenuation is mechanical. Panel (b): what remains after netting the benchmark out.

Table 8: Appendix A1: worst years for the constructed bond return

	Total return
2022	-0.164
2009	-0.102
2013	-0.085
1999	-0.078
1994	-0.072
1969	-0.056
2021	-0.041
1967	-0.032

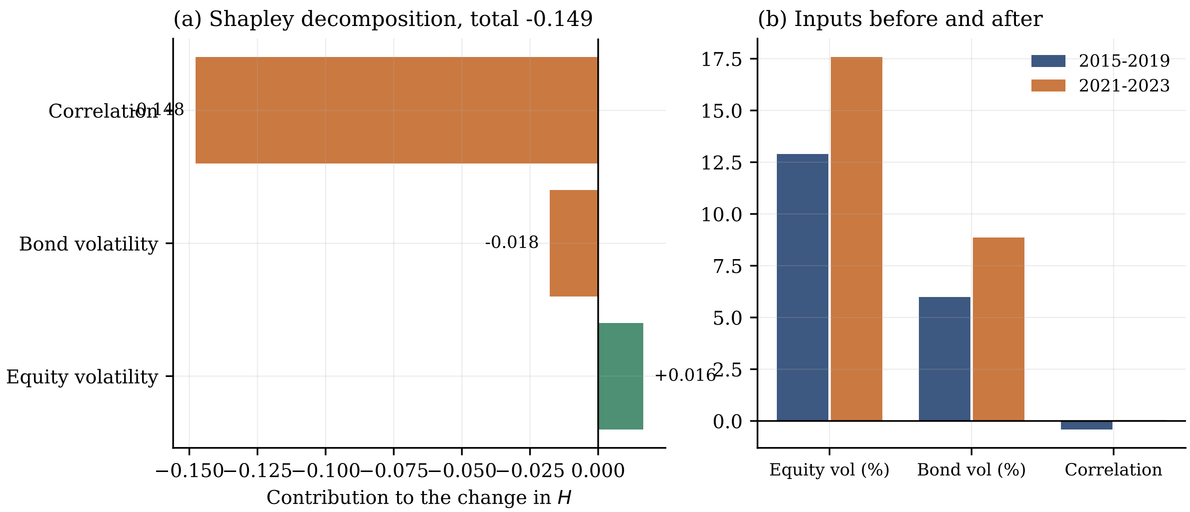

Figure 3: Decomposition of the change in the hedging benefit. Bond volatility rose, but so did equity volatility, and only their ratio enters.

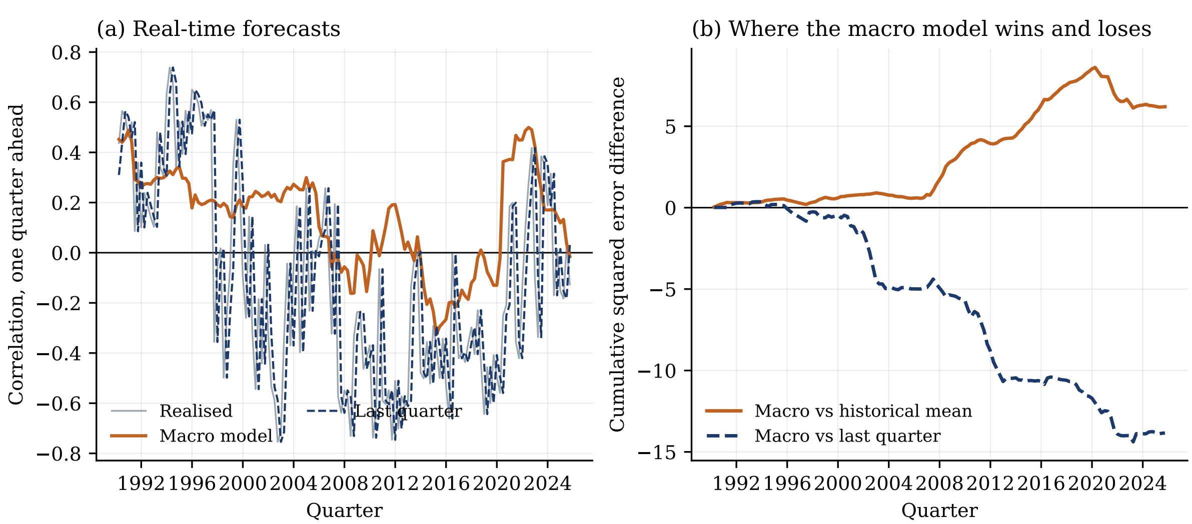

Figure 4: Real-time forecasts of the one-quarter-ahead correlation and the cumulative squared error difference against each benchmark.

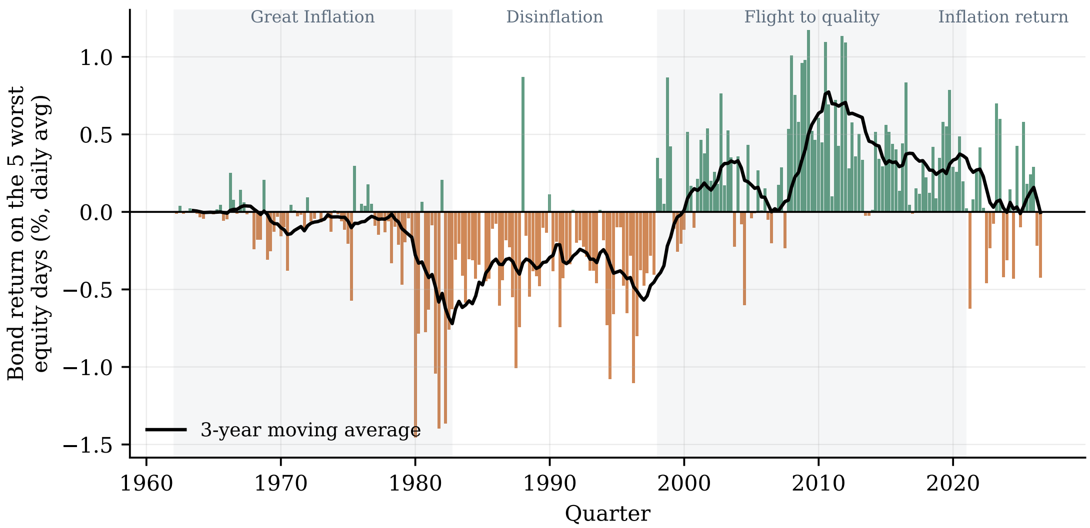

Figure 5: **The crash hedge weakened but did not invert.** Average daily bond excess return on the five worst equity days of each quarter.