HEDGE FUND STRATEGY PERFORMANCE ACROSS FINANCIAL AND GEOPOLITICAL CRISES: EVIDENCE FROM LIQUID REPLICATION INDICES

Luka Bulajić*
Singidunum University, Belgrade, Serbia

Abstract
This paper compares the crisis performance of five liquid hedge fund strategy indices, namely Long/Short Equity, Event Driven, Merger Arbitrage, Managed Futures and a multi-strategy Global Strategies index, together with their aggregate, against the S&P 500, a 60/40 portfolio and gold across five crises spanning 2008 to 2026. The crises differ in character, encompassing systemic financial, pandemic, geopolitical and policy-driven episodes. Each crisis is divided into a crisis phase and, for the four completed episodes, a recovery phase. The geopolitical episodes are examined alongside the daily Geopolitical Risk index, and differences in Sharpe ratios are tested using the Jobson-Korkie statistic with the Memmel correction and the Ledoit-Wolf bootstrap. The results show that crisis protection is regime-specific. Managed Futures hedged the systemic, pandemic and commodity-driven geopolitical crises but failed in the policy-driven tariff turmoil; Merger Arbitrage was the steadiest performer across the set, and gold protected capital in only two of the five episodes. Crisis phases range from 23 to 129 trading days, and many of the crisis-phase differences prove too small relative to their sampling variation to reach statistical significance. The underperformance of Managed Futures during recoveries is the most consistently significant result in the study, and two of its 2008 recovery comparisons are the only results to survive a Holm correction for multiple testing. The paper concludes that no single liquid alternative strategy can be relied upon as an unconditional crisis hedge, and that a combination of complementary strategies is more resilient than reliance on the strategy that performed best in the most recent episode.

Keywords
hedge fund strategies, crisis performance, liquid replication indices, managed futures, geopolitical risk, Sharpe ratio tests

JEL Classification
G01, G11, G14, G23

* Corresponding author, **Luka Bulajić** - luka.bulajic.251@singimail.rs

Introduction

Investors hold alternative strategies in large part for what they are expected to do when traditional markets fall. The diversifying appeal of hedge funds rests on a claim that their flexible mandates, use of derivatives and capacity to take short and directional positions allow them to preserve capital or even profit, while long-only portfolios suffer. Whether this promise survives contact with real crises, and whether it survives equally across crises of different character, is an empirical question with direct consequences for portfolio construction.

The period from 2008 to 2026 spans an unusually varied set of crises. Within it fall a systemic financial crisis (the 2008 Global Financial Crisis), a pandemic-driven liquidity shock (COVID-19 in early 2020), and a cluster of geopolitical and policy crises: the Russia-Ukraine war beginning in February 2022, the United States tariff turmoil of spring 2025, and the United States-Iran confrontation of 2026. These episodes differ not only in magnitude but in kind, in their speed, their persistence, and the asset markets through which they transmit. What protects a portfolio in one type of crisis may offer no protection at all in another.

The five episodes were selected so that together they cover the main types of crisis (systemic, pandemic, geopolitical and policy-driven), and so that the sample gives particular weight to the geopolitical and policy shocks of 2022-2026, which existing studies of hedge fund strategies have rarely examined. Two candidate episodes were considered and excluded. The 2011 euro area sovereign debt crisis was a second credit crisis, centred on European sovereign and banking markets; with United States benchmarks, and with the 2008 crisis already representing the systemic-credit type, it would duplicate a category rather than add one. The 2022 interest rate shock cannot be separated cleanly from the Russia-Ukraine episode, since the Federal Reserve began raising rates within weeks of the invasion and continued throughout both phases of that window. The Russia-Ukraine results are therefore best read as reflecting a combined geopolitical and monetary shock.

Most existing empirical evidence on hedge fund strategy behavior in crises is drawn from samples that end before this cluster of events, and almost none of it examines the investable, liquid replication versions of these strategies that are now available to allocators. Conventional hedge fund indices aggregate the returns that funds report voluntarily to commercial databases, and are exposed to the biases of that reporting, notably self-selection and the smoothing of reported returns (Fung and Hsieh, 2000; Getmansky, Lo and Makarov, 2004). Liquid replication indices are instead rules-based portfolios of liquid, tradable instruments, constructed to reproduce the return profile of a strategy category; they are priced daily, are free of reporting biases, and represent what an allocator seeking liquid exposure to a strategy can actually hold. This paper addresses both gaps. Five liquid hedge fund strategy indices and their aggregate are compared against conventional benchmarks across the five crises listed above. Each crisis is divided into a crisis phase and, for the four completed episodes, a recovery phase, and the geopolitical episodes are related to the Geopolitical Risk (GPR) index of Caldara and Iacoviello (2022).

The central question of this paper is whether crisis protection is regime-specific: whether any single strategy protects across all crisis types, or whether each type of crisis rewards

a different strategy. The evidence favours the latter. Managed Futures hedged the systemic, pandemic and commodity-driven geopolitical crises effectively, but reversed sharply in the policy-driven tariff crisis. Merger Arbitrage was the steadiest performer across the set, and gold protected capital in only two of the five episodes. These findings extend the cross-strategy crisis literature into the geopolitically dominated 2020-2026 period using liquid indices, and document the conditional, regime-dependent nature of strategy-level crisis protection.

The remainder of the paper is organized as follows. Section 1 reviews the relevant literature, Section 2 describes the data and method, and Section 3 presents and discusses the results. The final section concludes and considers the implications for portfolio construction.

1. Review of the scientific literature

The theoretical case for hedge funds as crisis diversifiers rests on their exposure to non-traditional sources of risk and return. Bali, Brown and Caglayan (2014) show that hedge funds' exposure to macroeconomic uncertainty explains a significant part of the cross-sectional dispersion in their returns, with the most exposed funds earning higher subsequent returns. Lhabitant (2006) and Mallaby (2010) provide the institutional and historical backdrop, documenting how leverage and derivative positioning have allowed certain managers to profit during episodes of market stress while also, as in the case of Long-Term Capital Management, occasionally amplifying it.

A second strand examines how crisis behaviour differs across strategies. Agarwal and Naik (2004) show that many hedge fund returns resemble the payoff of a short put option, embedding hidden tail risk that is realized precisely in down markets, and Agarwal, Ruenzi and Weigert (2017), using portfolio holdings, show that tail risk helps explain differences in the returns of equity-oriented hedge funds and stems largely from holdings of tail-sensitive stocks, option positions and leverage. Sadka (2010), using a broad sample of funds, finds that funds highly exposed to liquidity risk outperformed those with low exposure by about 6% a year between 1994 and 2008 but performed poorly during liquidity crises; critically for the present study, managed futures is the only style in which this spread was not sharply negative in the fourth quarter of 2008. Stoforos, Degiannakis and Palaskas (2017) reach a complementary conclusion: using dynamic correlations with the S&P 500, they reject the claim of crisis protection for most hedge fund strategies, the exceptions being macro and short-selling approaches.

The diversifying properties of trend-following and managed futures in particular have long been recognized. Fung and Hsieh (2001) show that the returns of trend followers resemble those of lookback straddles on bonds, currencies and commodities, so that they profit from large price moves in either direction, and Hurst, Ooi and Pedersen (2017) find that trend following has been profitable for more than a century and has tended to perform well during the largest drawdowns of a traditional 60/40 portfolio. Ang (2014) notes that, whereas most hedge fund strategies are in effect short volatility, trend-following strategies tend to gain when volatility rises and therefore tend to do well in bad times, while Gliner (2014) gives a practitioner account of global macro trading, in which managers take directional positions across asset classes in response to changing economic and policy conditions. Sullivan and Wey (2025) qualify this view, finding that the alpha of global

macro and managed futures funds has eroded since the 2008 financial crisis and that their downside protection has not consistently improved on that of a comparable stock-bond portfolio.

Because the present study uses liquid replication indices rather than reported fund returns, the replication literature is directly relevant. Hasanhodzic and Lo (2007) demonstrate that linear clones constructed from a small set of liquid factors can reproduce a meaningful portion of the return and risk profile of several hedge fund strategy categories. Their results are pertinent in two specific ways: managed futures replicates well and is structurally liquid, whereas event-driven and emerging-market strategies replicate poorly because the liquid clone cannot capture the illiquidity premium those strategies earn. Fung and Hsieh (2004) provide the seven-factor risk-based benchmarking framework underlying much of this work, and Getmansky, Lo and Makarov (2004) show that return smoothing, and hence the gap between reported and economic risk, is modest for the more liquid strategy categories, which supports the use of standard performance ratios for liquid indices.

The methodology used in this paper draws on a further body of work concerned with how performance differences should be tested rather than merely described. Jobson and Korkie (1981), as corrected by Memmel (2003), derive the standard error of the difference between two Sharpe ratios, accounting for the correlation between the return series being compared. Ledoit and Wolf (2008) develop a bootstrap alternative that relaxes the assumptions of normality and serial independence, and is therefore better suited to short samples of daily returns. Both procedures are applied here, for reasons set out in section 2.5.

Finally, a rapidly growing literature links geopolitical risk to asset prices. Caldara and Iacoviello (2022) construct the Geopolitical Risk index used here and show that higher geopolitical risk foreshadows lower investment and employment and is associated with larger downside risks. Rafi and Ali (2025) build separate return factors for geopolitical acts and geopolitical threats across 40 countries, finding that threats affect stock returns in more countries than acts, and Będowska-Sójka, Demir and Zaremba (2022) examine which assets hedged the Russia-Ukraine shock specifically. The present paper brings these strands together, asking how a panel of liquid hedge fund strategy indices behaved across a sequence of financial and geopolitical crises, and relating the geopolitical episodes to the GPR index.

2. Research methodology

2.1. Strategy indices and benchmarks

The strategy data are six daily liquid replication indices from the LAB Liquid family of HedgeIndex LLC, the provider of the hedge fund indices formerly published under the Credit Suisse name (HedgeIndex, 2026): five strategy indices, namely Long/Short Equity, Event Driven, Merger Arbitrage, Managed Futures and a multi-strategy index labelled Global Strategies, together with the LAB Liquid aggregate of the sub-indices. The index data were downloaded from the HedgeIndex platform (www.hedgeindex.com) as daily net asset values in USD, covering the period through 17 June 2026. The series were launched between May 2008, when the Long/Short Equity index went live, and 2011, when the Managed Futures index was introduced, and the former Global Macro index was broadened and renamed Global Strategies. Values before each launch date are

hypothetical performance data back-calculated by the provider, so that in the 2008 episode the Long/Short Equity index is the only one of the six indices that reflects live values throughout. Two points of interpretation are important and are maintained throughout. First, these are liquid, rules-based replication indices rather than reported returns of live funds; following Hasanhodzic and Lo (2007) and Fung and Hsieh (2004), liquid replication is treated as a transparent and investable proxy for strategy-level exposures, with the explicit caveat that it may understate both the returns and the crisis-period drawdowns of illiquidity-driven strategies. Second, the Global Strategies index is a multi-strategy basket rather than a global macro index, and is described as such throughout. The indices are nested: Merger Arbitrage is a component of the Event Driven index, Managed Futures is a component of the Global Strategies index, and the aggregate combines the Long/Short, Event Driven and Global Strategies indices. The set is a selection of accessible strategies rather than the full hedge fund universe.

Three benchmarks represent the traditional portfolio against which the strategies are judged: the S&P 500 (total-return basis), a 60/40 portfolio formed from 60% equities and 40% investment-grade bonds and rebalanced daily, and gold. Daily rebalancing holds the 60/40 weights constant within even the shortest windows, so that the benchmark reflects the allocation itself rather than the drift of its weights between rebalancing dates; with monthly or quarterly rebalancing instead, its crisis-phase returns change by less than one percentage point, and none of the comparisons reported below is affected. Benchmark price histories are obtained from Yahoo Finance using liquid exchange-traded funds on a dividend-adjusted (total-return) basis to match the net-asset-value basis of the strategy indices: SPY for the S&P 500, AGG for investment-grade bonds in the 60/40 portfolio, and GLD for gold. Because the LAB indices are published daily, including weekends, on which net asset values are carried forward unchanged, the strategy and benchmark series are intersected on common trading days; all statistics reported below are computed on that basis.

2.2. Geopolitical risk measure

Geopolitical conditions are captured by the daily Geopolitical Risk (GPR) index of Caldara and Iacoviello (2022), together with its decomposition into a geopolitical acts component and a geopolitical threats component, which is discussed in section 3.3 as a direction for further work. The GPR series were obtained from the index authors' data page (www.matteoiacoviello.com). Because the index measures geopolitical rather than financial or pandemic risk, it is overlaid only on the geopolitical and policy crises (Russia-Ukraine, the 2025 tariff episode and the 2026 United States-Iran confrontation) and not on the 2008 or COVID-19 crises. The index is elevated during the geopolitical episodes and low during the other two, which supports treating it as a measure of geopolitical rather than general market stress.

2.3. Crisis windows

Five crises are examined in this study. Each window is defined by identifiable market and event dates and, with one exception, is divided into a crisis phase (the drawdown) and a recovery phase. Crisis phases begin either at the market peak preceding the crisis or,

where the crisis was triggered by a specific event, at that event; they end at the market trough or at the point at which the acute phase of the crisis passed. Recovery phases run from the end of the crisis phase to the point at which the drawdown was recovered, or the crisis subsided. The 2026 United States-Iran crisis is treated as a crisis phase only, because at the end of the sample its resolution was incomplete. The windows and their classification are reported below (table no. 1).

Table no. 1. Crisis windows, types and phases

Crisis	Type	Crisis phase	Recovery phase
2008 GFC	Systemic / credit	02 Sep 2008 – 09 Mar 2009	10 Mar 2009 – 30 Sep 2009
COVID-19	Pandemic / liquidity	19 Feb 2020 – 23 Mar 2020	24 Mar 2020 – 31 Aug 2020
Russia-Ukraine	Geopolitical / commodity	24 Feb 2022 – 16 Jun 2022	17 Jun 2022 – 31 Dec 2022
US Tariffs 2025	Policy / trade	19 Feb 2025 – 08 Apr 2025	09 Apr 2025 – 27 Jun 2025
US-Iran 2026	Geopolitical / oil	04 Mar 2026 – 07 Apr 2026	ongoing

Notes: dates are inclusive. Window boundaries correspond to market peaks and troughs and to event dates, including the period immediately preceding the Lehman Brothers failure, the COVID-19 market peak, the start of the Russia-Ukraine war, the 2025 tariff announcements and Iran's declared closure of the Strait of Hormuz in early March 2026, which followed the United States and Israeli strikes of 28 February; the United States-Iran crisis phase ends on 7 April 2026, the day on which a two-week ceasefire was agreed (Congressional Research Service, 2026).

Source: author's classification, based on S&P 500 market data and publicly reported event dates.

Because crisis phases defined by market peaks and troughs are fixed with hindsight, and could favor strategies that move against equities, the main results are also re-estimated on fixed event-based windows that use no market turning points. Each such window opens at the close of the last trading day before a dated event and runs for a fixed 21 or 63 trading days in every crisis. The events are the Lehman Brothers bankruptcy filing on 15 September 2008, the declaration of the COVID-19 pandemic by the World Health Organization on 11 March 2020, the Russian invasion of Ukraine on 24 February 2022, the tariff announcement made after the close of trading on 2 April 2025, and the United States and Israeli strikes on Iran on 28 February 2026. Because the pandemic declaration came three weeks into the COVID-19 sell-off, a second COVID-19 window opening with the first Italian lockdowns of 21-23 February 2020 is also reported. The results are discussed in section 3.6.

2.4. Performance metrics

The analysis is conducted on daily data throughout. For each asset, in each crisis phase, the holding-period return, the maximum drawdown and the Sharpe ratio are computed, together with each asset's correlation with the S&P 500 within the window. The Sharpe ratio uses a zero risk-free rate, 252 trading days per year, and the total-return convention of Hasanhodzic and Lo (2007); the modest smoothing documented by Getmansky, Lo and Makarov (2004) for liquid strategies justifies the use of standard ratios. Because

short-term interest rates averaged below one percent in the 2008, COVID-19 and Russia-Ukraine crisis phases but above three percent in the later windows, all Sharpe ratios and tests are also recomputed on returns in excess of the three-month United States Treasury bill rate (section 3.6). For the geopolitical crises, the thirty-day rolling correlation of each strategy with the S&P 500 is plotted against the GPR index.

Two limitations of the method are stated plainly. The analysis is a calendar-window event study across a small number of episodes rather than a large-sample exercise; several windows are short, and the formal tests described in section 2.5 are therefore reported alongside the descriptive statistics so that the strength of each comparison can be judged directly. And, as noted, the liquid indices may understate both the calm-period returns and the crisis-period drawdowns of illiquidity-driven strategies.

2.5. Statistical inference

Because several crisis windows are short, apparent differences in performance may lie within the bounds of sampling error: the COVID-19 and United States-Iran crisis phases each contain only 23 common trading days and the 2025 tariff phase 34, against 129 for the 2008 crisis phase and 78 for Russia-Ukraine. Differences in risk-adjusted performance are therefore tested formally. For each pair of assets in each phase, the null hypothesis is that the two annualized Sharpe ratios are equal, and it is assessed in two ways. The first is the closed-form statistic of Jobson and Korkie (1981) as corrected by Memmel (2003), which accounts for the correlation between the two return series but assumes normally distributed and serially independent returns. The second is a studentized block bootstrap following the approach of Ledoit and Wolf (2008), who show that inference on Sharpe ratio differences should allow for heavy tails and serial dependence in returns. Returns are resampled in overlapping circular blocks of length equal to the cube root of the number of observations, the bootstrap statistic is studentized using the standard error of Memmel (2003), and a two-sided p-value is obtained from 4,999 resamples. The bootstrap is treated here as the primary test, with the closed-form statistic reported alongside as a conventional benchmark.

Testing is restricted to the comparisons on which the argument rests: each strategy against the S&P 500, and Managed Futures against each other strategy. This gives eleven comparisons in each of the nine phase-windows, or 99 tests in total. Because about five of 99 tests would be expected to reject at the 5% level by chance alone, the p-values are also adjusted using the step-down procedure of Holm (1979), which controls the probability of any false rejection across the full set of tests. Two results survive the adjustment under both tests: the underperformance of Managed Futures relative to Global Strategies and to the LAB aggregate in the 2008 recovery phase. Unadjusted p-values are reported in the discussion that follows, since the comparisons were specified in advance by the argument rather than selected after inspecting the results, but individual findings at the 5% level should be read with the adjustment in mind. The number of observations underlying each test is reported so that statistical power can be judged directly. The full set of computed performance metrics and the complete results of the pairwise tests are provided in the online appendix.

3. Results and discussion

The crisis-phase holding-period return of every asset in every crisis is reported below (table no. 2). The first observation is that all six liquid indices, the five strategies and their aggregate, lost less than the S&P 500 in every crisis: where the S&P fell 46.0% in 2008, 33.7% during the COVID-19 crash, and between 14.1% and 18.8% in the Russia-Ukraine and tariff episodes, no strategy index lost more than 17.6%. Against the more defensive 60/40 portfolio, the picture is less uniform, and in the 2025 tariff crisis three of the six indices fell further than the balanced benchmark. In capital-preservation terms, the diversification promise holds against equities, but not unconditionally. On a risk-adjusted basis, moreover, most of these differences are not statistically significant, as section 3.5 shows: the sampling distribution of Sharpe differences is wide. The more interesting result lies in the dispersion across strategies, which is large and, crucially, changes sign from one crisis to the next.

Table no. 2. Crisis-phase holding-period returns (%)

Asset	2008 GFC	COVID-19	Russia-Ukr.	Tariffs 2025	US-Iran 2026
Long/Short Equity	−12.4	−12.2	−5.6	−12.0	+0.2
Global Strategies	−14.5	−6.7	−0.4	−13.5	+0.4
LAB Liquid (Aggregate)	−10.4	−9.7	−3.0	−12.4	+0.2
Event Driven	−14.1	−17.6	−9.8	−7.8	−0.8
Merger Arbitrage	−8.7	−6.8	−4.2	−0.2	+1.2
Managed Futures	+17.1	+7.6	+15.3	−11.1	+0.8
S&P 500	−46.0	−33.7	−14.1	−18.8	−3.5
60/40	−29.0	−21.7	−11.4	−11.3	−2.5
Gold	+14.4	−3.6	−2.5	+1.6	−8.5

Notes: holding-period return over each crisis phase as defined in table no. 1. The upper six rows are the liquid indices, five strategies and their aggregate; the lower three rows are benchmarks.

Source: author's own research.

The corresponding crisis-phase Sharpe ratios, correlations with the S&P 500 and maximum drawdowns are reported below (table no. 3).

Table no. 3. Crisis-phase Sharpe ratios, correlations with the S&P 500 and maximum drawdowns

Asset	2008 GFC	COVID-19	Russia-Ukr.	Tariffs 2025	US-Iran 2026
Panel A. Sharpe ratio					
Long/Short Equity	−1.18	−6.26	−1.63	−5.47	0.31
Global Strategies	−2.49	−7.75	−0.23	−6.54	1.13
LAB Liquid (Aggregate)	−1.80	−7.84	−1.30	−6.23	0.58
Event Driven	−2.18	−9.91	−2.06	−5.17	−1.57
Merger Arbitrage	−1.15	−2.83	−2.51	−0.27	4.93
Managed Futures	2.03	3.33	3.28	−5.20	0.75
S&P 500	−1.82	−5.46	−1.66	−5.61	−2.14
60/40	−1.70	−5.27	−2.16	−5.59	−2.23

Gold	0.85	−1.05	−0.41	0.84	−2.60
Panel B. Correlation with the S&P 500					
Long/Short Equity	0.21	0.98	0.72	0.94	−0.17
Global Strategies	0.62	0.92	0.74	0.93	0.60
LAB Liquid (Aggregate)	0.51	0.95	0.83	0.95	0.46
Event Driven	0.67	0.87	0.89	0.92	0.95
Merger Arbitrage	0.71	0.65	0.68	0.59	0.24
Managed Futures	−0.47	−0.60	−0.35	0.48	−0.50
S&P 500	1.00	1.00	1.00	1.00	1.00
60/40	0.98	0.98	0.98	0.99	0.99
Gold	0.04	−0.01	−0.10	0.27	0.42
Panel C. Maximum drawdown (%)					
Long/Short Equity	−20.2	−12.2	−9.0	−12.0	−2.1
Global Strategies	−15.4	−6.9	−3.4	−13.5	−0.9
LAB Liquid (Aggregate)	−16.1	−9.7	−5.7	−12.4	−1.3
Event Driven	−17.3	−17.6	−12.1	−7.8	−2.2
Merger Arbitrage	−14.8	−11.2	−5.3	−2.2	−0.3
Managed Futures	−8.0	−4.3	−3.9	−11.7	−3.2
S&P 500	−46.0	−33.7	−20.6	−18.8	−7.5
60/40	−29.0	−21.7	−14.7	−11.3	−5.1
Gold	−22.1	−12.5	−12.0	−5.0	−16.2

Notes: Sharpe ratios are annualized with a zero risk-free rate; correlations are computed from daily returns; maximum drawdowns are measured from the running peak within each crisis phase as defined in table no. 1.
Source: author's own research.

3.1. Managed Futures: the conditional hedge

Managed Futures is the strategy whose behavior most sharply illustrates regime-specificity. In the 2008 systemic crisis, it returned +17.1%, in the COVID-19 crash +7.6%, and during the Russia-Ukraine crisis phase +15.3%, in every case while equities fell heavily and with a maximum drawdown of 8% or less. Its crisis-phase Sharpe ratios in these three episodes ranged from 2.0 to 3.3, and its correlation with the S&P 500 was negative throughout, for example −0.47 in 2008 and −0.60 during COVID-19 (table no. 3). The advantage over equities is significant at the 5% level in the COVID-19 crisis under the bootstrap test ($p = 0.04$), though not under the closed-form statistic ($p = 0.14$), and is marginal in 2008 and Russia-Ukraine ($p \approx 0.10$ under both tests); the detail is reported in table no. 4. This is the textbook profile of a trend-following crisis hedge: sustained, directional moves in rates, currencies and commodities are precisely what such strategies are built to capture, and the returns recorded in these three episodes are consistent with that mechanism.

The 2025 tariff crisis broke the pattern. There, Managed Futures lost 11.1%, a sharp reversal from its performance in the other crises, with a crisis-phase Sharpe of −5.2. The reversal is statistically visible in the comparison with Merger Arbitrage, the best performer of that window, which Managed Futures underperformed significantly ($p = 0.02$ under the bootstrap test); against equities, by contrast, the difference is negligible (p

= 0.83), reflecting that Managed Futures did not so much fall further than the market as fail to protect against it.

The reason is likely to lie in the character of the price action rather than its direction: the tariff episode consisted of abrupt announcements and partial reversals, a pattern that repeatedly breaks and re-establishes trends in the currency and commodity markets on which trend strategies rely. Index-level data do not permit attribution of the loss to individual asset classes, and identifying the specific markets responsible would require position-level or sector-level return data. This remains an open question for further research. The very mechanism that makes Managed Futures a hedge against durable crises makes it vulnerable to policy-driven volatility. The contrast is visible in figures no. 1 to no. 3: the strategy rides above the S&P 500 and the 60/40 portfolio for almost the whole of the 2008 and Russia-Ukraine windows, but tracks them downward in 2025.

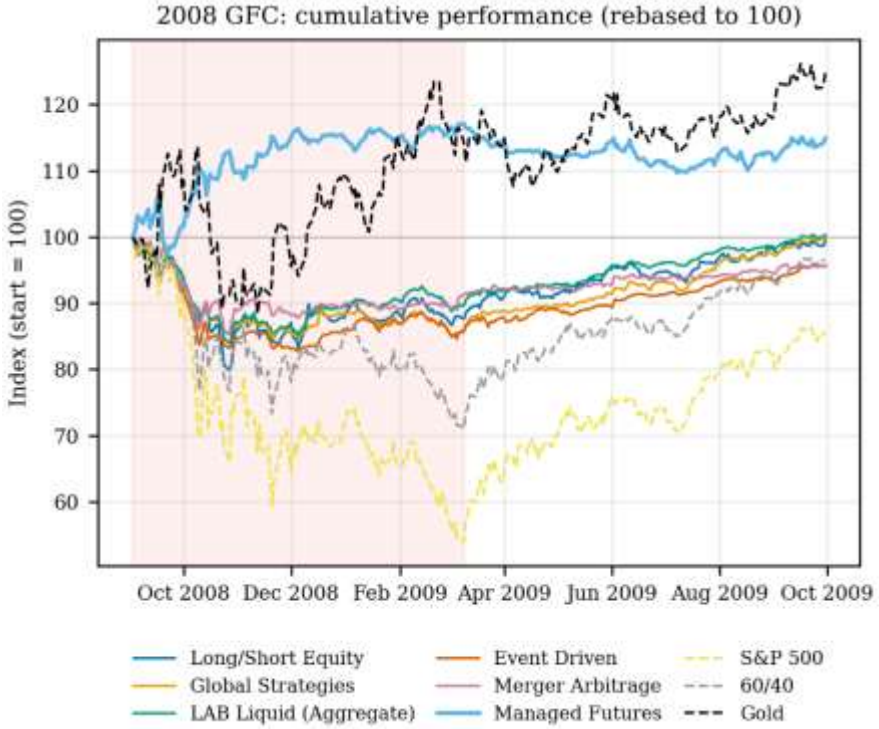

Figure no. 1: Cumulative performance during the 2008 Global Financial Crisis, all assets rebased to 100 at the window start; the shaded band marks the crisis phase
Source: author's own research.



Figure no. 2: Cumulative performance during the Russia-Ukraine crisis; the shaded band marks the crisis phase
Source: author's own research.

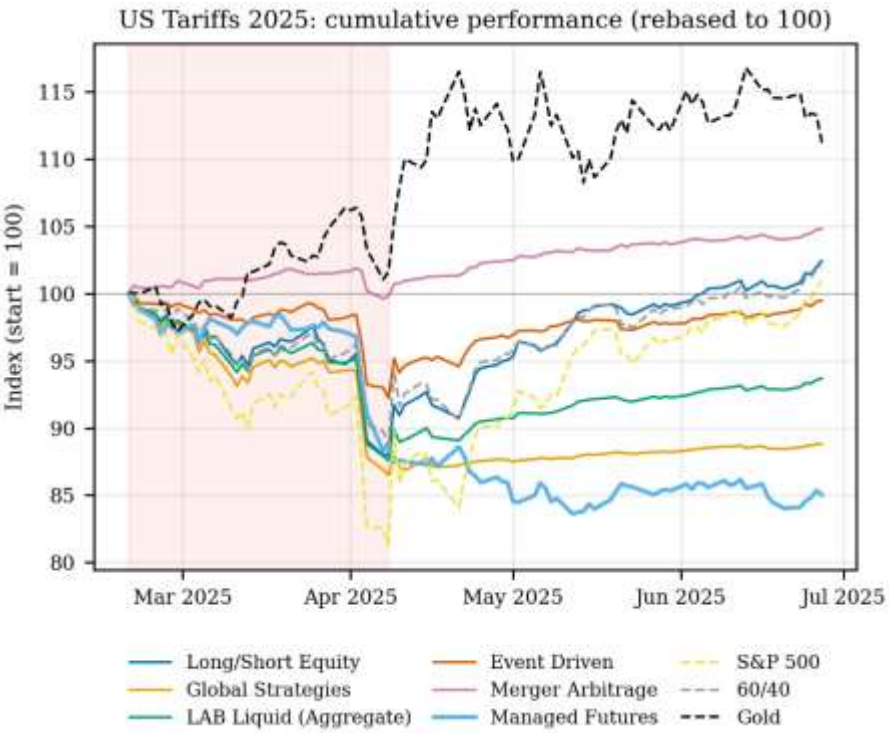

Figure no. 3: Cumulative performance during the 2025 tariff crisis; the shaded band marks the crisis phase
Source: author's own research.

3.2. The rotation of the best crisis hedge

Because no strategy wins everywhere, the identity of the best crisis hedge rotates with the type of crisis. In the systemic, pandemic and commodity-driven episodes it is Managed Futures; in the policy-driven tariff crisis the most resilient assets are Merger Arbitrage, which lost only 0.2%, and gold, which gained 1.6%; in the short, oil-driven United States-Iran episode the strategies were broadly flat to mildly positive, with Merger Arbitrage again best at +1.2% while gold fell 8.5%. The United States-Iran window is, however, too short to support inference: with only 23 common trading days, no comparison between strategies comes close to significance, the smallest p-value among them being 0.32. Only the gap between Merger Arbitrage and the S&P 500 clears the 5% threshold, and then only under the bootstrap test ($p = 0.03$ against $p = 0.09$ for the closed-form statistic). That episode is therefore described here, but no weight is placed on the ranking within it. This rotation is the core finding: crisis protection is conditional on the nature of the crisis, and an allocator who selected a single crisis hedge on the basis of one episode would have been poorly served in the next.

Two secondary patterns reinforce the point. Gold, often described as a reliable haven, protected capital in only two of the five crises, 2008 and the 2025 tariff episode, and fell

during COVID-19, Russia-Ukraine and the United States-Iran episode, the last by 8.5%. Its protective value is therefore selective rather than general. Merger Arbitrage, by contrast, was the most consistent performer of the set: its worst crisis-phase loss across the entire sample was 8.7%, in 2008, and it was the single best strategy in both the tariff and United States-Iran crises. Formal testing supports this in the tariff window, where its advantage over both equities ($p = 0.048$) and Managed Futures ($p = 0.02$) is significant, and in the tariff recovery, where it outperformed Managed Futures with $p = 0.001$. Its advantage over equities also reaches significance in the Russia-Ukraine recovery ($p = 0.04$) and, under the bootstrap alone, in the short United States-Iran window; elsewhere its margins fall within sampling error. These three advantages over equities, in the tariff crisis, the Russia-Ukraine recovery and the United States-Iran window, do not survive the use of a Treasury bill risk-free rate, whereas its advantage over Managed Futures in the tariff window does (section 3.6). Its low and relatively stable correlation with equities, rather than any dramatic negative beta, is what delivered this steadiness. These findings align with Sadka (2010), in whose sample managed futures was the only style in which funds more exposed to liquidity risk did not sharply underperform those less exposed in the fourth quarter of 2008, and are consistent with Stoforos, Degiannakis and Palaskas (2017), who rejected the claim of crisis protection for most hedge fund strategies but not for macro and short-selling approaches, which, like managed futures, can take directional short positions; the present results qualify both by identifying the policy-driven regime in which the trend hedge fails.

3.3. Correlation behaviour and geopolitical risk

For the geopolitical crises, the rolling correlation of each strategy with the S&P 500, set against the GPR index, clarifies the mechanism of the Managed Futures hedge. The Russia-Ukraine episode is shown below (figure no. 4): as the GPR index spikes, the Managed Futures correlation with the equity market falls well below zero, into the region of −0.5 to −0.85, while the other strategies remain positively correlated throughout. The hedge is therefore not a matter of low correlation in general but of active negative correlation precisely when geopolitical risk is elevated. Rafi and Ali (2025) find that geopolitical threats and geopolitical acts affect stock returns differently. Separating the response to these two components offers a natural refinement of the present analysis, which is left for further work. The fact that the same GPR index stays low during the 2008 and COVID-19 crises, which were not geopolitical in origin, supports its use as a specifically geopolitical, rather than generic, stress indicator.

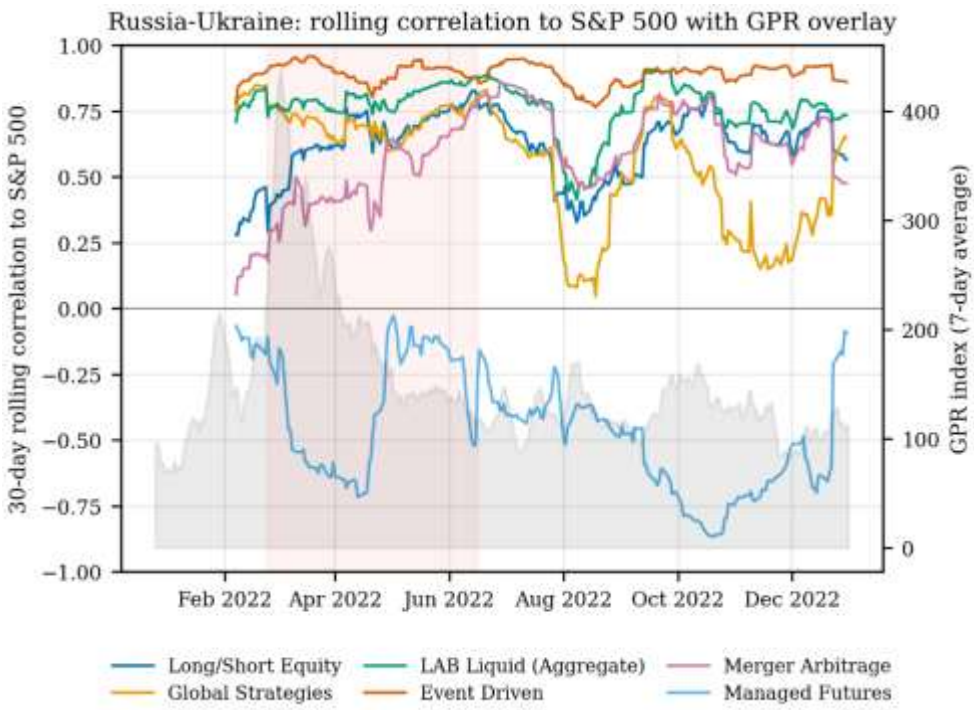

Figure no. 4: Thirty-day rolling correlation of each strategy with the S&P 500 (left axis) and the GPR index (right axis, shaded) during the Russia-Ukraine crisis; the red shaded band marks the crisis phase
Source: author's own research.

3.4. Drawdowns and the recovery trade-off

The drawdown perspective sharpens the comparison. In 2008, shown below (figure no. 5), the S&P 500 fell to a trough drawdown of 46% and the 60/40 portfolio to 29%, whereas the maximum drawdown of Managed Futures was held to about 8% and those of the other strategy indices ranged from 14.8% to 20.2% (table no. 3, panel C). These maximum drawdowns, measured from the running peak within the crisis phase, are larger than the phase losses of 8.7% to 14.5% reported in table no. 2, because the strategy indices had recovered part of their losses by the end of the phase on 9 March 2009, whereas the S&P 500 and the 60/40 portfolio ended it at their trough. Measured by the size of the loss, the downside protection offered by the strategies is real and substantial.

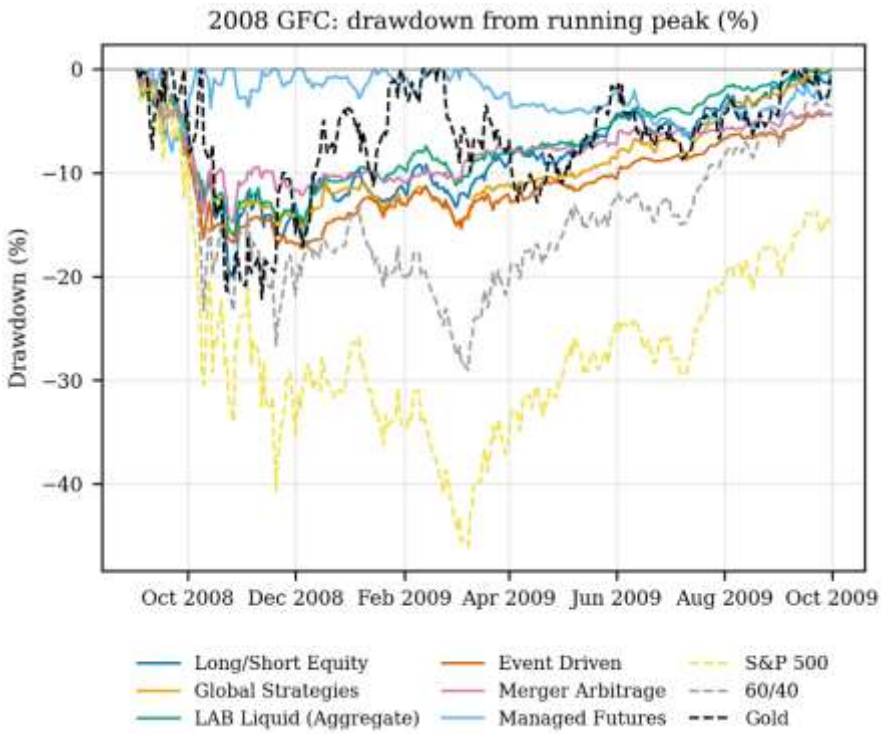

Figure no. 5: Drawdown from the running peak during the 2008 crisis
Source: author's own research.

The recovery phases reveal the cost of that protection. The strategies that defended best in the drawdown generally lagged in the rebound, and Managed Futures in particular lost ground in every recovery phase studied, with negative recovery-phase returns following each of the 2008, COVID-19, Russia-Ukraine and tariff crises. This is also the most statistically robust result in the study. In the 2008 recovery, Managed Futures underperformed the multi-strategy Global Strategies index and the LAB aggregate at better than the 0.1% level, and Long/Short Equity, Event Driven and the S&P 500 at the 5% level under the bootstrap test; the pattern repeats in the COVID-19 recovery and, most sharply, in the tariff recovery, where it lagged Merger Arbitrage with $p = 0.001$. The two 2008 comparisons with Global Strategies and the LAB aggregate are also the only results in the study to survive the Holm adjustment for multiple testing (section 2.5). Where the crisis-phase advantages of trend following are often too noisy to establish formally, the recovery-phase shortfall is not. The protection Managed Futures offers on the way down is thus paid for, in part, by underperformance on the way up, a trade-off that a return measured only over the drawdown would conceal, and one that justifies the two-phase design adopted here. For an allocator, the practical implication is that the value of a trend hedge depends not only on whether it limits the loss but on how its full-cycle behavior fits the rest of the portfolio.

The end of the 2008 recovery phase on 30 September 2009 requires comment, since the S&P 500 had not recovered its drawdown by that date and did not regain its pre-crisis peak for several years. A recovery phase running to full recovery would have extended far into the post-crisis expansion, where returns reflect the subsequent economic cycle rather than the crisis itself. Under the rule set out in section 2.3, the phase therefore ends at the point at which the crisis subsided: roughly six months after the March 2009 trough, when the rebound in equities was well established, and the stress in funding markets had largely eased. The Russia-Ukraine recovery phase, which ends with the close of 2022 although the war continued, is defined on the same basis, whereas the COVID-19 and tariff recovery phases end once the S&P 500 had recovered from its drawdown.

3.5. Statistical significance of the results

The formal tests for the comparisons on which the discussion rests are collected below (table no. 4). Three patterns stand out. First, the descriptive advantages recorded in table no. 2 are frequently not significant on a risk-adjusted basis: in the 2008 crisis phase, the largest and most visually striking of the episodes, no strategy-versus-equity comparison rejects the null of equal Sharpe ratios at the 5% level. Second, significance depends far more on the size of the gap than on the length of the window: the two longest crisis phases, 2008 and Russia-Ukraine, produce no significant strategy-versus-equity results at all, while the 23-day COVID-19 phase does, because the differences there are much larger. In the equally short United States-Iran phase, significance is almost absent: the only significant result, Merger Arbitrage against the S&P 500, holds under the bootstrap alone. Third, the two tests disagree in 11 of the 99 comparisons computed, most of them in the shorter windows, and the disagreement is systematic: the bootstrap rejects more often than the closed-form statistic, which is the expected consequence of daily returns that are neither normal nor serially independent. Where the two disagree, the bootstrap result is preferred, and the disagreement is reported rather than resolved silently.

Table no. 4. Tests of differences in annualized Sharpe ratios (selected comparisons)

Comparison	Crisis	Phase	N	Δ Sharpe	p (JKM)	p (LW)
Managed Futures vs S&P 500	2008 GFC	crisis	129	3.85	0.109	0.098
Managed Futures vs S&P 500	COVID-19	crisis	23	8.79	0.144	0.044
Managed Futures vs S&P 500	Russia-Ukraine	crisis	78	4.93	0.097	0.098
Managed Futures vs S&P 500	US Tariffs 2025	crisis	34	0.41	0.887	0.834
Managed Futures vs S&P 500	US-Iran 2026	crisis	23	2.89	0.615	0.565
Merger Arbitrage vs S&P 500	2008 GFC	crisis	129	0.67	0.533	0.529
Merger Arbitrage vs S&P 500	COVID-19	crisis	23	2.63	0.356	0.493
Merger Arbitrage vs S&P 500	Russia-Ukraine	crisis	78	−0.85	0.559	0.517
Merger Arbitrage vs S&P 500	US Tariffs 2025	crisis	34	5.34	0.037	0.048

Merger Arbitrage vs S&P 500	US-Iran 2026	crisis	23	7.07	0.090	0.029
Managed Futures vs Merger Arbitrage	US Tariffs 2025	crisis	34	−4.93	0.086	0.018
Managed Futures vs Merger Arbitrage	US Tariffs 2025	recovery	54	−9.69	0.001	0.001
Managed Futures vs S&P 500	2008 GFC	recovery	142	−3.28	0.050	0.013
Managed Futures vs S&P 500	COVID-19	recovery	111	−4.30	0.094	0.065

Notes: Δ Sharpe is the annualized Sharpe ratio of the first asset named less that of the second, so a negative value indicates that the first underperformed. The null hypothesis in each test is that the two Sharpe ratios are equal, and the reported p-values are two-sided. p (JKM) is the Jobson-Korkie statistic with the Memmel (2003) correction; p (LW) is the studentized block bootstrap described in section 2.5. N is the number of common trading days in the window. Values below 0.05 are treated as significant and values below 0.10 as marginal.
Source: author's own research.

Throughout, the stated caveats apply. The sample comprises five episodes and nine phase-windows, and, as table no. 4 makes clear, many of the descriptive differences are too noisy to establish formally; the patterns are economically clear, but only a subset of them is statistically significant. The liquid indices may also understate the crisis-period losses of the strategies they replicate. In the case of Merger Arbitrage in particular, a rules-based replication need not fully capture the deal-break and financing risk that a live portfolio of announced transactions carries into a crisis, so its consistency here should be read as a property of the investable, liquid form rather than as a claim about live funds pursuing the strategy. With these qualifications, the evidence consistently supports a single conclusion: strategy-level crisis protection is regime-specific.

3.6. Robustness checks

Three checks address the main choices of method. The first replaces the crisis phases of table no. 1, whose boundaries are set by market peaks and troughs, with the fixed event-based windows described in section 2.3; the holding-period returns are reported below (table no. 5). The central pattern for Managed Futures survives. It gained after the Lehman Brothers failure (7.5% over 21 trading days and 12.5% over 63) and after the Russian invasion of Ukraine (9.2% and 13.3%), and lost after the 2025 tariff announcement (12.7% and 12.0%), where its shortfall against Merger Arbitrage is significant under both tests ($p = 0.002$ over 63 trading days). In the United States-Iran windows, it was flat to negative, consistent with the absence of any significant result for Managed Futures in that episode in the main analysis. Merger Arbitrage again avoided large losses, its worst being 8.8% over 21 trading days and 9.9% over 63, both in 2008, and in the 21 trading days after the strikes on Iran it gained 1.0% while the S&P 500 fell 7.6%, a difference significant under both tests ($p = 0.008$ and $p = 0.001$).

Table no. 5. Holding-period returns in fixed event-based windows (%)

Asset	2008 GFC	COVID-19 (WHO)	COVID-19 (Italy)	Russia-Ukr.	Tariffs 2025	US-Iran 2026
Panel A. 21 trading days from the event						
Long/Short Equity	−8.8	−1.3	−11.8	+0.6	+1.0	+0.3

Global Strategies	−8.6	−2.8	−6.7	+1.9	−7.0	−1.4
LAB Liquid (Aggregate)	−8.6	−2.9	−9.6	+1.6	−4.2	−1.0
Event Driven	−11.8	−6.4	−17.5	+2.5	−1.2	−2.0
Merger Arbitrage	−8.8	−2.4	−6.8	+0.1	+0.9	+1.0
Managed Futures	+7.5	−4.0	+7.1	+9.2	−12.7	−2.9
S&P 500	−19.2	−4.4	−32.8	+7.1	+0.4	−7.6
60/40	−14.8	−2.1	−21.2	+3.4	+0.3	−5.4
Gold	+8.5	+0.1	−5.4	+2.7	+3.4	−14.3
Panel B. 63 trading days from the event						
Long/Short Equity	−7.6	+8.6	−0.7	−4.1	+7.7	−2.4
Global Strategies	−11.7	+0.5	−3.8	+0.1	−5.6	+1.4
LAB Liquid (Aggregate)	−10.0	+2.8	−3.6	−1.9	−1.1	+0.4
Event Driven	−15.5	+2.5	−7.9	−6.8	+1.8	+1.6
Merger Arbitrage	−9.9	−1.8	−3.5	−2.0	+3.7	+2.3
Managed Futures	+12.5	−7.8	−1.6	+13.3	−12.0	0.0
S&P 500	−29.9	+11.9	−11.1	−6.4	+11.1	+10.6
60/40	−17.6	+8.4	−4.7	−5.5	+7.1	+5.7
Gold	+6.8	+4.4	+4.9	−2.3	+6.6	−13.8

Notes: each window opens at the close of the last common trading day before the event (12 September 2008; 10 March 2020 for the pandemic declaration and 21 February 2020 for the Italian lockdowns; 23 February 2022; 2 April 2025, the day of the after-hours tariff announcement; and 27 February 2026) and runs for 21 or 63 common trading days. The events are described in section 2.3.

Source: author's own research.

Two results depend on the window. The first is the COVID-19 protection offered by Managed Futures, which depends on the date from which the crisis is measured. In the window opening with the Italian lockdowns, it gained 7.1% over 21 trading days while the S&P 500 fell 32.8%, but in the window opening with the pandemic declaration of 11 March 2020 it lost 4.0% over 21 trading days and 7.8% over 63, as the sharp reversal of markets after 23 March cut across the trends it had followed into the sell-off. The trend hedge protects against a sustained decline but is exposed when that decline reverses abruptly, the same mechanism that undermined it in the tariff episode. The second is that the longer windows mix the crisis with the rebound that followed: by the end of the 63 trading days opening with the pandemic declaration, the tariff announcement and the strikes on Iran, the S&P 500 had gained 11.9%, 11.1% and 10.6% respectively, so that these windows no longer isolate the period of stress. The Russia-Ukraine windows are instructive in a different way, since the S&P 500 rose 7.1% in the 21 trading days after the invasion and fell only later, as monetary tightening took hold, which supports reading that episode as a combined geopolitical and monetary shock. Fixed windows therefore complement, rather than replace, the two-phase design adopted here.

The second check replaces the zero risk-free rate with the three-month United States Treasury bill rate. The rate averaged below one per cent in the 2008 and COVID-19 windows and in the Russia-Ukraine crisis phase, where no Sharpe ratio changes by more than 0.16, and between 3.2% and 4.2% in the later windows. The two comparisons that

survive the Holm adjustment are unaffected. Under the bootstrap test, the number of comparisons significant at the 5% level falls from 23 to 20, and the three that lose significance are all comparisons of Merger Arbitrage with the S&P 500, in the Russia-Ukraine recovery, the tariff crisis and the United States-Iran crisis. Under the closed-form statistic, the count is unchanged at 16, as the Merger Arbitrage comparisons with equities in the tariff crisis and recovery lose significance, and two comparisons whose p-values lay just above the threshold (0.050) fall just below it. The explanation is mechanical. Merger Arbitrage was far less volatile than equities in these windows, with annualized volatility of 2.2% to 5.6% against 17.6% to 26.7% for the S&P 500, so that deducting a Treasury bill rate of more than three percent lowers its Sharpe ratio much more than that of equities. Part of its advantage under a zero risk-free rate therefore reflects a return close to that of cash rather than outperformance of the equity market.

The third check concerns the 60/40 benchmark. Rebalancing it monthly or quarterly instead of daily changes its crisis-phase returns by less than one percentage point, the largest difference arising in 2008 (a loss of 29.9% with monthly rebalancing against 29.0%), and alters none of the comparisons reported above.

Conclusions

This paper compared five liquid hedge fund strategy indices and their aggregate with conventional benchmarks across five crises spanning 2008 to 2026, dividing each crisis into a crisis phase and, for the four completed episodes, a recovery phase and relating the geopolitical episodes to the Geopolitical Risk index. Two findings stand out. First, in capital-preservation terms, every strategy index lost less than equities in every crisis, and most also lost less than a balanced 60/40 portfolio, which supports the diversification value of these investable strategies in their liquid, replicated form; on a risk-adjusted basis, however, most of these differences are not statistically significant, and the claim is therefore descriptive rather than established. Second, and more importantly, the identity of the best-performing strategy rotated with the character of the crisis: Managed Futures was a powerful hedge in the systemic, pandemic and commodity-driven geopolitical crises but failed in the policy-driven tariff turmoil; Merger Arbitrage was the steadiest performer across the set; and gold protected in only two of the five episodes. The central pattern for Managed Futures survives re-estimation on fixed event-based windows and with a Treasury bill risk-free rate, with two qualifications: the COVID-19 protection offered by Managed Futures depends on the date from which the crisis is measured, and part of the advantage of Merger Arbitrage over equities in the high-rate episodes of 2022 to 2026 reflects a return close to that of cash. Of these results, the most statistically robust is not a crisis-phase finding at all but the recovery-phase underperformance of Managed Futures, which is significant across several episodes and includes the only two comparisons to survive a correction for multiple testing.

The practical implication for allocators is that no single liquid alternative strategy can be relied upon as an unconditional crisis hedge. The protection a strategy offers is conditional on the mechanism of the crisis, whether systemic, pandemic, geopolitical-commodity or policy-driven. Holding several complementary strategies is therefore more robust than relying on the one that happened to perform best in the most recent episode.

The study is limited by its small number of episodes, by the fact that many comparisons cannot be established formally, by its reliance on liquid replication indices, and, in the 2008 episode, by the use of hypothetical, back-calculated values for all indices except Long/Short Equity. Because such indices are marked continuously against liquid, tradable instruments, they understate not only the returns of the strategies they replicate but also their crisis-period losses: live funds holding infrequently valued positions, or restricting redemptions under stress, are likely to have experienced deeper drawdowns than those recorded here (Getmansky, Lo and Makarov, 2004). Future research could extend the panel to additional strategy categories where investable replications exist, incorporate transaction costs explicitly, attribute crisis-period losses to individual asset classes using position-level data, extend formal inference to differences in returns and drawdowns as well as Sharpe ratios, and revisit the regime-specificity result as the sample of crises grows and statistical power improves. The broad message, however, is unlikely to be overturned: in a period defined increasingly by geopolitical and policy shocks, the question is not whether a strategy hedges crises, but which crises it hedges.

References

[1] Agarwal, V. and Naik, N.Y., 2004. Risks and portfolio decisions involving hedge funds. *Review of Financial Studies*, 17(1), pp.63-98.

[2] Agarwal, V., Ruenzi, S. and Weigert, F., 2017. Tail risk in hedge funds: a unique view from portfolio holdings. *Journal of Financial Economics*, 125(3), pp.610-636.

[3] Ang, A., 2014. *Asset management: a systematic approach to factor investing*. New York: Oxford University Press.

[4] Bali, T.G., Brown, S.J. and Caglayan, M.O., 2014. Macroeconomic risk and hedge fund returns. *Journal of Financial Economics*, 114(1), pp.1-19.

[5] Będowska-Sójka, B., Demir, E. and Zaremba, A., 2022. Hedging geopolitical risks with different asset classes: a focus on the Russian invasion of Ukraine. *Finance Research Letters*, 50, 103192.

[6] Caldara, D. and Iacoviello, M., 2022. Measuring geopolitical risk. *American Economic Review*, 112(4), pp.1194-1225.

[7] Congressional Research Service, 2026. *The Strait of Hormuz: security developments and impacts on oil, gas, and other commodities*. CRS Report R45281, updated 7 August 2026. Washington, DC: Congressional Research Service.

[8] Fung, W. and Hsieh, D.A., 2000. Performance characteristics of hedge funds and commodity funds: natural vs. spurious biases. *Journal of Financial and Quantitative Analysis*, 35(3), pp.291-307.

[9] Fung, W. and Hsieh, D.A., 2001. The risk in hedge fund strategies: theory and evidence from trend followers. *Review of Financial Studies*, 14(2), pp.313-341.

[10] Fung, W. and Hsieh, D.A., 2004. Hedge fund benchmarks: a risk-based approach. *Financial Analysts Journal*, 60(5), pp.65-80.

[11] Getmansky, M., Lo, A.W. and Makarov, I., 2004. An econometric model of serial correlation and illiquidity in hedge fund returns. *Journal of Financial Economics*, 74(3), pp.529-609.

[12] Gliner, G., 2014. *Global macro trading: profiting in a new world economy*. Hoboken: Bloomberg Press/Wiley.

[13] Hasanhodzic, J. and Lo, A.W., 2007. Can hedge-fund returns be replicated?: the linear case. *Journal of Investment Management*, 5(2), pp.5-45.

[14] HedgeIndex, 2026. *LAB Liquid Indexes*. [online] Available at: <https://www.hedgeindex.com> [Accessed 11 November 2026].

[15] Holm, S., 1979. A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6(2), pp.65-70.

[16] Hurst, B., Ooi, Y.H. and Pedersen, L.H., 2017. A century of evidence on trend-following investing. *Journal of Portfolio Management*, 44(1), pp.15-29.

[17] Jobson, J.D. and Korkie, B.M., 1981. Performance hypothesis testing with the Sharpe and Treynor measures. *Journal of Finance*, 36(4), pp.889-908.

[18] Ledoit, O. and Wolf, M., 2008. Robust performance hypothesis testing with the Sharpe ratio. *Journal of Empirical Finance*, 15(5), pp.850-859.

[19] Lhabitant, F.-S., 2006. *Handbook of hedge funds*. Chichester: John Wiley and Sons.

[20] Mallaby, S., 2010. *More money than God: hedge funds and the making of a new elite*. New York: Penguin Press.

[21] Memmel, C., 2003. Performance hypothesis testing with the Sharpe ratio. *Finance Letters*, 1, pp.21-23.

[22] Rafi, M.K.H. and Ali, S.R.M., 2025. Disaggregated geopolitical risks and global stock returns. *Global Finance Journal*, 67, 101151.

[23] Sadka, R., 2010. Liquidity risk and the cross-section of hedge-fund returns. *Journal of Financial Economics*, 98(1), pp.54-71.

[24] Stoforos, C.E., Degiannakis, S. and Palaskas, T.B., 2017. Hedge fund returns under crisis scenarios: a holistic approach. *Research in International Business and Finance*, 42, pp.1196-1207.

[25] Sullivan, R.N. and Wey, M., 2025. Global macro and managed futures hedge fund strategies: portfolio differentiators? *Journal of Alternative Investments*, 27(4), pp.41-65.