Finance Research Letters 106 (2026) 110194

Contents lists available at ScienceDirect

Finance Research Letters

journal homepage: www.elsevier.com/locate/frl

# Gold, Bitcoin, and equity market stress: Evidence from Double Machine Learning

Marcelo J. Villena [a,*], Mauricio G. Villena [b]

[a] Departamento de Ingenieria Comercial, Universidad Técnica Federico Santa María, Chile
[b] Departamento de Administración, Facultad de Administración y Economía, Universidad Diego Portales, Chile

ARTICLE INFO

*JEL classification:*
G11
G15
C14
C55

*Keywords:*
Gold
Bitcoin
Safe-haven assets
Equity market stress
Local projections
Double machine learning
VIX

ABSTRACT

We reexamine whether Gold and Bitcoin protect U.S. equity portfolios during market stress. Using daily data from December 2014 to February 2026, we combine a contemporaneous Baur–Lucey safe-haven benchmark, horizon-specific local projections, and an orthogonalised Double Machine Learning (DML) stress-interaction estimator. The integrated design separates three questions that the literature usually conflates: whether an asset hedges equities on average, whether it protects on the day of a stress realisation, and whether any protection persists over the subsequent holding period. The headline finding is for Bitcoin: the contemporaneous stress interaction with the S&P 500 is $\delta_0 = 0.124$ with $p \approx 1.1 \times 10^{-6}$, positive (the opposite of the safe-haven prediction) and Bonferroni-significant at the family-wise 5% level. Under VIX-defined stress, the orthogonalised DML stress interaction for Bitcoin is positive and Bonferroni-significant at three of five horizons; the more demanding nonlinear conditioning strategy reinforces rather than overturns the rejection. Gold's stress interactions are weak, unstable, and imprecisely estimated; we do not reject a defensive role for Gold but the data do not support a robust safe-haven role either. A multi-allocation portfolio grid and a spot-Gold robustness exercise corroborate the regression evidence. Economically, the results are more consistent with liquidity-driven liquidation and risk-on/risk-off dynamics than with either asset as reliable equity crash insurance.

## 1. Introduction

Whether Gold and Bitcoin provide effective protection against equity-market risk remains an open question. Gold has long been viewed as a defensive asset during financial turbulence, while Bitcoin has increasingly been promoted as "digital gold". Yet the evidence remains mixed, varying across sample periods, stress definitions, horizons, and empirical methods.

This paper reexamines the relation between Gold, Bitcoin, and U.S. equity risk along two complementary dimensions: (i) whether *contemporaneous* innovations in Gold and Bitcoin returns offset same-day equity outcomes, and (ii) whether those innovations are associated with future S&P 500 returns and losses over short and medium horizons. The first is the canonical safe-haven test in the spirit of Baur and Lucey (2010), Baur and McDermott (2010): does the asset move in the right direction at the moment of the equity shock? The second matters for an investor who holds Gold or Bitcoin as a hedge over a non-trivial holding period, since even on-impact protection is only useful if it persists.

* Corresponding author.
*E-mail address:* marcelo.villena@usm.cl (M.J. Villena).

https://doi.org/10.1016/j.frl.2026.110194
Received 8 March 2026; Received in revised form 3 May 2026; Accepted 15 May 2026
Available online 4 June 2026
1544-6123/© 2026 Elsevier Inc. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

To address (i) we report a contemporaneous Baur–Lucey-style benchmark. To address (ii) we estimate horizon-specific local projections under two designs: a standard linear local projection (OLS-LP) and a Double Machine Learning local projection (DML-LP) in which the conditioning information set enters flexibly through machine-learning-based nuisance estimation. The role of DML is deliberately modest: it is a robustness device, not a methodological contribution. Following Baur and Lucey (2010), Baur and McDermott (2010), a safe haven is an asset that becomes especially protective in adverse states. We therefore examine stress-state interactions under two definitions of stress: a VIX-based high-volatility state, and a tail-event definition based on the lower tail of daily S&P 500 returns.

Recent contributions sharpen the conditional view further. Chibane and Janson (2025) document that Bitcoin and the Swiss franc – not Gold – function as safe havens against geopolitical equity-crash risk, while Gurdgiev and Petrovskiy (2024) directly question whether Gold is always a safe haven. Tarchella et al. (2024) and Echaust et al. (2024) reach broadly compatible conclusions for cryptocurrencies as inconsistent equity hedges across the pre- and post-COVID periods. We contribute to this conditional-safe-haven literature by integrating the contemporaneous Baur–Lucey test, multi-horizon local projections, and an orthogonalised DML stress-interaction estimator within a common sample.

Using daily data from December 2014 to February 2026, the results can be summarised as follows. **The headline finding is for Bitcoin.** The contemporaneous Baur–Lucey stress interaction with the S&P 500 is $\delta_0 = 0.124$ with $p \approx 1.1 \times 10^{-6}$, surviving Bonferroni correction at the family-wise 5% level. On the day of a stress realisation, Bitcoin moves *with* the equity market, not against it. Under the orthogonalised DML stress-interaction estimator, Bitcoin's VIX-stress interaction is positive and Bonferroni-significant at three of five horizons ($h = 0, 1, 5$); the more demanding nonlinear conditioning strengthens rather than weakens the rejection. **Gold is different.** Its stress-state interactions are positive in sign at most horizons but uniformly imprecise: no horizon clears the Bonferroni-adjusted threshold. We therefore do not claim the data *reject* a defensive role for Gold, but they do not support a robust safe-haven role either. **Neither method nor horizon choice explains the negative result:** the contemporaneous and multi-horizon designs deliver the same qualitative answer, and OLS-LP and DML-LP agree closely at the short horizons that carry the headline finding. **The portfolio and robustness evidence corroborates the regression results**—a multi-allocation grid varying Gold and Bitcoin weights from 0% to 20% confirms Gold as a weak, capacity-limited diversifier and Bitcoin as a return-enhancing but drawdown-amplifying risk asset; a spot-Gold robustness exercise using continuous COMEX Gold futures preserves the sign pattern of all GLD-based coefficients.

Our value added is to separate three questions the literature usually conflates – hedging on average, on-impact protection during stress, and persistence over the holding period – and to adopt an asymmetric position the literature rarely makes carefully: Bitcoin clearly fails the safe-haven test at multiple-testing-corrected significance, while the data are too imprecise to reject a defensive role for Gold.

The remainder of the paper proceeds as follows. Section 2 reviews the related literature. Section 3 presents the empirical framework. Section 4 describes the data and implementation. Section 5 reports the main results. Section 6 concludes.

## 2. Related literature

This paper relates to three strands of literature.

First, on Gold as a hedge and safe haven, Baur and Lucey (2010), Baur and McDermott (2010) establish the distinction between average hedging and protection during severe market stress, and later work emphasises that Gold's defensive role is conditional rather than uniform. Reboredo (2013), O'Connor et al. (2015), He et al. (2018), and Bampinas and Panagiotidis (2017) characterise Gold as a "sometime safe haven" whose role is highly sensitive to sample, market structure, and design. Bredin et al. (2017) extend this conditional view from a downside-risk perspective, closely aligned with the loss-based specification we use below. More recently, Gurdgiev and Petrovskiy (2024) directly question whether Gold is always a safe haven.

Second, on Bitcoin as a potential hedge or safe haven, early contributions (Dyhrberg, 2016a,b) drew analogies between Bitcoin and Gold, but later work became more cautious. Bouri et al. (2017), Baur et al. (2018), Klein et al. (2018), and Smales (2019) find Bitcoin to be an unstable substitute for Gold, while Shahzad et al. (2020) continues to favour Gold. Crisis-period studies (Conlon and McGee, 2020; Conlon et al., 2020) show that cryptocurrencies often moved with risky assets during COVID-19. Syuhada et al. (2022) reach a related conclusion using a vine-copula approach. Tarchella et al. (2024) and Echaust et al. (2024) document inconsistent hedge properties across pre- and post-COVID periods, and Chibane and Janson (2025) report that Bitcoin protects against geopolitical equity-crash risk while Gold does not. Gkillas and Longin (2019) and Selmi et al. (2018) reinforce the view that Bitcoin's defensive role is limited and unstable.

Third, on empirical methodology, Jordà (2005) provides the local-projection framework, and we draw on the orthogonalised DML logic of Chernozhukov et al. (2018) as a pragmatic robustness device.

Relative to this literature, we reexamine whether Gold and Bitcoin provide robust protection against U.S. equity risk in a common sample, over common horizons, and under explicit stress-state tests, using both contemporaneous and forward-looking designs. The evidence is more consistent with unstable or procyclical comovement than with a persistent defensive role.

## 3. Empirical framework

### 3.1. Contemporaneous benchmark

Let $R_t^{SP}$ denote the date-$t$ return of the S&P 500 and $R_t^H$ the return on a candidate hedge $H \in \{\text{Gold}, \text{Bitcoin}\}$. The Baur–Lucey contemporaneous test is

$$R_t^{SP} = \alpha + \theta_0 R_t^H + \delta_0 \left(R_t^H \times S_t\right) + \gamma' Z_t + u_t, \tag{1}$$

where $S_t$ is a stress indicator and $Z_t$ collects controls. A negative $\theta_0$ is consistent with average hedging; a negative $\delta_0$ is consistent with safe-haven behaviour during stress.

### 3.2. Local projections

For each horizon $h \in \{0, 1, \ldots, H\}$ we estimate

$$Y_{t,h} = \theta_h D_t + g_h(X_t) + u_{t+h}, \tag{2}$$

where $Y_{t,h} \equiv \sum_{j=0}^{h} R_{t+j}^{SP}$ and $D_t \equiv R_t^H$. The vector $X_t$ collects macro-financial controls (lagged returns of the S&P 500, Gold, and Bitcoin, the VIX, the U.S. Dollar Index, Treasury yields, credit spreads, and the term spread). Horizon $h = 0$ reduces to Eq. (1) without the stress interaction, so the multi-horizon design extends rather than replaces the contemporaneous benchmark. A negative $\theta_h$ is consistent with hedging.

Because hedging is often interpreted as protection against losses, we also estimate

$$L_{t,h} = \phi_h D_t + q_h(X_t) + e_{t+h}, \quad L_{t,h} \equiv -\sum_{j=0}^{h} R_{t+j}^{SP}, \tag{3}$$

with $\phi_h > 0$ consistent with protection against future equity losses.

The benchmark estimator is a linear local projection,

$$Y_{t,h} = \alpha_h + \theta_h D_t + \Gamma_h' X_t + u_{t+h}, \tag{4}$$

estimated by OLS with Newey–West standard errors. For robustness, we compare this with a Double Machine Learning local projection (DML-LP). Define the nuisance functions $\ell_h(X_t) = \mathbb{E}[Y_{t,h} \mid X_t]$ and $m(X_t) = \mathbb{E}[D_t \mid X_t]$, estimated using Random Forests with chronological cross-fitting. The target coefficient is recovered from

$$\widetilde{Y}_{t,h} = \theta_h \widetilde{D}_t + \varepsilon_{t,h}, \tag{5}$$

where $\widetilde{Y}_{t,h} = Y_{t,h} - \widehat{\ell}_h(X_t)$ and $\widetilde{D}_t = D_t - \widehat{m}(X_t)$. Close agreement between OLS-LP and DML-LP is informative in itself: it indicates that the weak hedge performance is not an artefact of restrictive linear controls.[1]

### 3.3. Stress-state interaction with orthogonalised DML

To evaluate whether an asset becomes more protective in stress periods, we estimate

$$Y_{t,h} = \theta_h D_t + \delta_h (D_t \times S_t) + g_h(X_t) + u_{t+h}. \tag{6}$$

In the DML-LP version, we treat $T_t^{(1)} = D_t$ and $T_t^{(2)} = D_t \times S_t$ as two regressors of interest, residualise each on $\{S_t, X_t\}$ via cross-fitted Random Forests, residualise the outcome on the same set, and recover the coefficients from

$$\widetilde{Y}_{t,h} = \theta_h \widetilde{T}_t^{(1)} + \delta_h \widetilde{T}_t^{(2)} + \varepsilon_{t,h}. \tag{7}$$

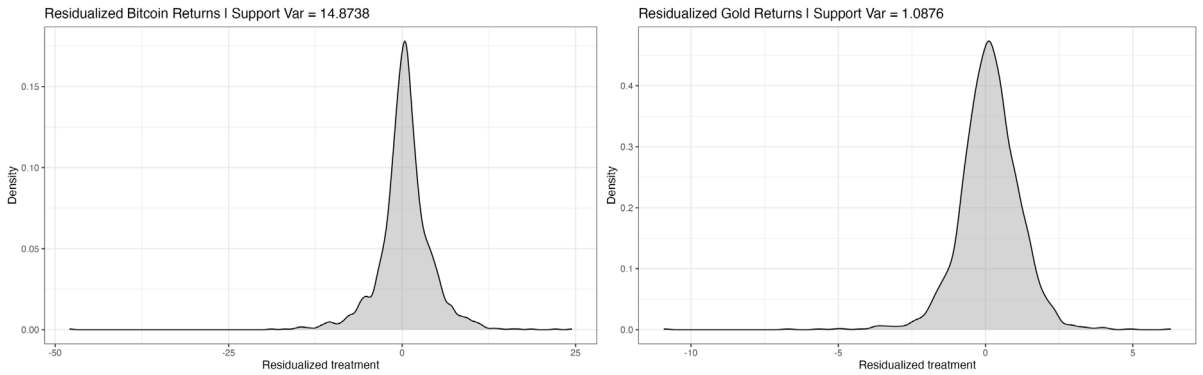
This residualises the interaction term in the same way as the baseline treatment, yielding a more defensible orthogonal moment than sequential partialling-out.

We use two stress definitions: a VIX-based state ($S_t = 1$ when the VIX is in its upper sample quartile) and a tail-event state (S&P 500 daily return in the lower 5% tail; we additionally report a 1% tail as a robustness check).

## 4. Data and implementation

The empirical analysis uses daily data on the S&P 500, Gold, Bitcoin, the VIX, the U.S. Dollar Index, and Treasury yields. The common sample runs from December 2, 2014 to February 27, 2026 and contains 2160 daily observations. Bitcoin is measured in U.S. dollars. Returns are computed as log differences in percentage terms.

[1] Re-estimating with $K = 5$ and $K = 10$ chronological folds confirms that varying cross-fitting folds does not meaningfully alter coefficient magnitudes, signs, or significance (Appendix A).




(a) Bitcoin residualised treatment (b) Gold residualised treatment

**Fig. 1.** Residualised treatment densities in the DML specification.
*Notes:* Kernel density estimates of $\widetilde{D}_t = D_t - \hat{m}(X_t)$, the residualised candidate hedge return after Random-Forest cross-fitted partialling-out of macro-financial controls $X_t$. Sample: daily data, December 2, 2014 to February 27, 2026, $N = 2160$. Wider residual support indicates a wider region of independent variation available for identification. Bitcoin displays substantially wider residual support than Gold (residual variances 14.8738 vs. 1.0876).

*Choice of gold proxy.* Gold is measured in the baseline using the GLD ETF, since it is the asset a U.S. equity investor actually holds to hedge equity drawdowns and trades on the same exchange and during the same hours as the S&P 500. To address ETF-specific frictions (creation-redemption flows, embeddedness in the U.S. equity complex), we replicate the baseline LP, DML-LP, downside, and VIX-stress specifications using the continuous COMEX Gold futures contract (Yahoo Finance ticker GC=F) as a spot-Gold proxy. Reported in Appendix B (Table 9), the signs of all coefficients match the GLD-based estimates; magnitudes are systematically smaller, consistent with ETF frictions amplifying procyclical comovement at high frequency. The qualitative conclusion is preserved, so we do not interpret the GLD-based results as an artefact of the ETF wrapper.

The information set includes lagged own returns, lagged returns of the other assets, lagged changes in the dollar index, lagged Treasury-yield changes, corporate credit spreads, and the term spread. We use a parsimonious lag structure based on the previous 12 daily observations. Dynamic effects are estimated at horizons $h \in \{0, 1, 5, 10, 20\}$ trading days. We also report rolling-window DML-LP estimates using windows of about 252 trading days. In the baseline return-based specification, the residual support variance is 14.8738 for Bitcoin and 1.0876 for Gold; both assets remain identified in the full sample, but Bitcoin is estimated from a wider region of independent variation (Fig. 1). A brief Monte Carlo exercise, reported in Appendix A, provides supporting evidence on finite-sample precision at longer horizons.

## 5. Results

### 5.1. Contemporaneous safe-haven benchmark

We begin with the canonical Baur–Lucey contemporaneous regression in Eq. (1), which addresses the concern that a safe-haven test should evaluate same-day protection rather than dynamic responses. Table 1 reports the contemporaneous estimates for Bitcoin and Gold under both VIX-based and tail-event definitions of stress.

The qualitative pattern is the same as in the dynamic local-projection design. The baseline contemporaneous coefficient $\theta_0$ is positive for Bitcoin, not negative as a hedge would require; it is also positive (and small) for Gold. The stress interaction $\delta_0$ is positive and significant for Bitcoin under VIX-based stress: $\delta_0 = 0.124$ with $p \approx 1.1\times10^{-6}$, which survives a Bonferroni correction at the family-wise 5% level. This is exactly the opposite of the safe-haven prediction. For Gold, $\delta_0$ is small and statistically indistinguishable from zero under both stress definitions. The contemporaneous benchmark therefore confirms that the LP-based message in Sections 5.2–5.3 is not a by-product of looking at future returns.

### 5.2. Average dynamic effects: benchmark versus flexible conditioning

Tables 2 and 3 report the main full-sample local-projection estimates. Across both specifications, neither Gold nor Bitcoin exhibits the sign pattern expected from a robust hedge against U.S. equity risk.

In the return-based specification, hedge-consistent evidence would require negative coefficients. Instead, both OLS-LP and DML-LP produce positive coefficients for Bitcoin at all horizons. Gold is also positive throughout, with larger effects at medium and long horizons. The DML estimate for Gold at the 20-day horizon is 0.237, implying that a one-percentage-point increase in Gold return today is associated with roughly a 0.24 percentage-point increase in cumulative S&P 500 return over the next 20 trading days. This is the opposite of what a robust hedge would imply.

**Table 1**
Contemporaneous safe-haven benchmark (Baur–Lucey style)

| | VIX-based stress | | Tail-event stress (5%) | |
|---|---|---|---|---|
| | Bitcoin | Gold | Bitcoin | Gold |
| Baseline ($\theta_0$) | 0.020$^{***B}$ (0.005) | 0.039 (0.025) | 0.031$^{***B}$ (0.006) | 0.072** (0.030) |
| Stress interaction ($\delta_0$) | 0.124$^{***B}$ (0.025) | 0.108 (0.082) | 0.046 (0.044) | 0.131 (0.166) |

*Notes:* Estimates from Eq. (1): contemporaneous OLS regression of the date-$t$ S&P 500 return on the date-$t$ candidate hedge return ($\theta_0$), the candidate hedge return interacted with a stress indicator ($\delta_0$), and macro-financial controls. Sample: daily data, December 2, 2014 to February 27, 2026, $N = 2160$. VIX-based stress: $S_t = 1$ when the VIX is in its upper sample quartile. Tail-event stress: $S_t = 1$ when the date-$t$ S&P 500 return is in the lower 5% tail of the sample. Newey–West standard errors in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively. The superscript $^B$ marks coefficients that survive a Bonferroni correction at the family-wise 5% level (with five horizon-specific tests, $\alpha/5 = 0.01$). Hedge-consistent evidence requires $\theta_0 < 0$; safe-haven-consistent evidence requires $\delta_0 < 0$. The Bitcoin ×VIX stress interaction $\delta_0 = 0.124$ is positive (the opposite of the safe-haven prediction) and survives Bonferroni correction with $p \approx 1.1 \times 10^{-6}$.

**Table 2**
Return-based dynamic effects: OLS-LP versus DML-LP.

| Horizon | Bitcoin OLS-LP | Bitcoin DML-LP | Gold OLS-LP | Gold DML-LP |
|---|---|---|---|---|
| 0 | 0.061*** (0.013) | 0.070*** (0.014) | 0.078** (0.039) | 0.092** (0.043) |
| 1 | 0.057*** (0.008) | 0.064*** (0.009) | 0.088** (0.044) | 0.107** (0.051) |
| 5 | 0.077*** (0.015) | 0.080*** (0.016) | 0.059 (0.057) | 0.079 (0.061) |
| 10 | 0.063*** (0.019) | 0.058*** (0.019) | 0.128 (0.079) | 0.188** (0.083) |
| 20 | 0.063*** (0.022) | 0.047* (0.025) | 0.208** (0.092) | 0.237* (0.127) |

*Notes:* Estimates from Eq. (2) via OLS-LP (Eq. (4)) and DML-LP (Eq. (5)). Dependent variable: cumulative future S&P 500 return $Y_{t,h} = \sum_{j=0}^{h} R_{t+j}^{SP}$, in percentage points. Horizon $h$ measured in trading days. Sample: daily data, December 2, 2014 to February 27, 2026, $N = 2160$. DML-LP nuisance functions estimated using Random Forests with $K = 2$ chronological folds. Negative coefficients are consistent with hedging. Newey–West standard errors in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

**Table 3**
Downside-based dynamic effects: OLS-LP versus DML-LP.

| Horizon | Bitcoin OLS-LP | Bitcoin DML-LP | Gold OLS-LP | Gold DML-LP |
|---|---|---|---|---|
| 0 | −0.061*** (0.013) | −0.069*** (0.014) | −0.078** (0.039) | −0.092** (0.043) |
| 1 | −0.057*** (0.008) | −0.064*** (0.009) | −0.088** (0.044) | −0.104** (0.051) |
| 5 | −0.077*** (0.015) | −0.081*** (0.016) | −0.059 (0.057) | −0.080 (0.061) |
| 10 | −0.063*** (0.019) | −0.058*** (0.019) | −0.128 (0.079) | −0.176** (0.082) |
| 20 | −0.063*** (0.022) | −0.048* (0.025) | −0.208** (0.092) | −0.224* (0.121) |

*Notes:* Estimates from Eq. (3). Dependent variable: cumulative future S&P 500 loss $L_{t,h} = -\sum_{j=0}^{h} R_{t+j}^{SP}$, in percentage points. Sample: daily data, December 2, 2014 to February 27, 2026, $N = 2160$. Positive coefficients are consistent with protection against future equity losses. Newey–West standard errors in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

The comparison between OLS-LP and DML-LP is informative in itself. At short horizons the two estimators agree closely, which tells us that the weak hedge performance is not an artefact of the linear specification. At medium and long horizons the DML conditioning changes some Gold magnitudes materially (e.g., 0.128 OLS vs. 0.188 DML at $h = 10$), but in no case does DML flip sign. Even under a substantially more flexible conditioning strategy the data continue to refuse a hedge interpretation—the principal economic value-added of the DML exercise.

The downside-based specification leads to the same conclusion from the opposite direction. If either asset protected against future equity losses, the coefficients should be positive. Instead, they are negative throughout. Taken together, the return-based and downside-based estimates point to procyclical rather than defensive average comovement.

### 5.3. Stress-state interaction results

The safe-haven question is more demanding than the hedge question because it asks whether an asset becomes especially protective in bad states. We evaluate stress dependence under VIX-based and tail-event definitions using the orthogonalised DML stress-interaction estimator in Eq. (7).

The VIX-based results in Table 4 are the strongest piece of evidence in the paper. For Bitcoin, the orthogonal DML stress interaction is positive and *Bonferroni-significant* at horizons $h = 0, 1$, and 5 (Panel B). At impact, the additional effect during stress is $\delta_0 = 0.129$ with a Newey–West standard error of 0.025 and a $p$-value of $2.4 \times 10^{-7}$, which clears the Bonferroni-adjusted 1% threshold within the panel. At one and five days the corresponding interactions are 0.066 ($p = 0.0015$) and 0.101 ($p = 0.00067$), again Bonferroni-significant. Rather than becoming more negatively related to future equity returns during stress, Bitcoin becomes more positively associated with them, and the result survives both the most demanding nonlinear conditioning strategy and a multiple-testing adjustment. We view this as the paper's headline empirical finding.

**Table 4**
Stress-state interaction estimates under VIX-based stress.

| Horizon | Bitcoin OLS-LP | Bitcoin DML-LP* | Gold OLS-LP | Gold DML-LP* |
|---|---|---|---|---|
| *Panel A. Baseline coefficient on asset return ($\theta_h$)* | | | | |
| 0 | 0.018*** (0.005) | 0.025***$^B$ (0.005) | 0.041 (0.028) | 0.069***$^B$ (0.023) |
| 1 | 0.030*** (0.007) | 0.040***$^B$ (0.008) | 0.041 (0.041) | 0.082***$^B$ (0.031) |
| 5 | 0.046*** (0.013) | 0.048***$^B$ (0.013) | 0.037 (0.061) | 0.073 (0.057) |
| 10 | 0.033* (0.017) | 0.026 (0.020) | −0.013 (0.091) | 0.090 (0.086) |
| 20 | 0.038* (0.022) | 0.030 (0.030) | 0.126 (0.110) | 0.170 (0.145) |
| *Panel B. Additional effect during stress ($\delta_h$)* | | | | |
| 0 | 0.126*** (0.024) | 0.129***$^B$ (0.025) | 0.096 (0.095) | 0.079 (0.096) |
| 1 | 0.077*** (0.022) | 0.066***$^B$ (0.021) | 0.122 (0.101) | 0.070 (0.106) |
| 5 | 0.090*** (0.031) | 0.101***$^B$ (0.030) | 0.056 (0.129) | 0.096 (0.142) |
| 10 | 0.085* (0.046) | 0.082* (0.045) | 0.368* (0.202) | 0.252 (0.187) |
| 20 | 0.074 (0.047) | 0.049 (0.056) | 0.214 (0.214) | 0.074 (0.290) |

*Notes:* Estimates from the stress-state interaction model in Eq. (6). Dependent variable: cumulative future S&P 500 return $Y_{t,h}$. VIX-based stress: $S_t = 1$ when the VIX is in its upper sample quartile. Sample: daily data, December 2, 2014 to February 27, 2026, $N = 2160$. *The DML-LP columns use the orthogonalised stress-interaction estimator in Eq. (7), which residualises both $D_t$ and $D_t \times S_t$ on $\{S_t, X_t\}$ via cross-fitted Random Forests. In the return-based specification, a negative interaction coefficient ($\delta_h < 0$) is consistent with stronger hedge behaviour during stress. Newey–West standard errors in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively. The superscript $^B$ marks coefficients that survive a Bonferroni correction across the five horizon-specific tests within each panel ($\alpha/5 = 0.01$). The Bitcoin DML-LP stress interaction is positive and Bonferroni-significant at horizons $h = 0, 1, 5$ (Panel B), the opposite of the safe-haven prediction.

**Table 5**
Stress-state interaction estimates under tail-event stress (5%)

| Horizon | Bitcoin OLS-LP | Bitcoin DML-LP* | Gold OLS-LP | Gold DML-LP* |
|---|---|---|---|---|
| *Panel A. Baseline coefficient on asset return ($\theta_h$)* | | | | |
| 0 | 0.029*** (0.006) | 0.032*** (0.006) | 0.054* (0.032) | 0.077** (0.034) |
| 1 | 0.044*** (0.008) | 0.051*** (0.009) | 0.059 (0.043) | 0.096** (0.042) |
| 5 | 0.057*** (0.013) | 0.064*** (0.015) | 0.026 (0.057) | 0.091 (0.059) |
| 10 | 0.047** (0.019) | 0.047** (0.021) | 0.086 (0.082) | 0.160** (0.080) |
| 20 | 0.061*** (0.022) | 0.048 (0.029) | 0.204** (0.090) | 0.163 (0.119) |
| *Panel B. Additional effect during stress ($\delta_h$)* | | | | |
| 0 | 0.049 (0.047) | 0.072 (0.044) | 0.194 (0.189) | 0.084 (0.203) |
| 1 | −0.081*** (0.031) | −0.024 (0.030) | 0.241 (0.147) | 0.147 (0.157) |
| 5 | −0.005 (0.053) | 0.048 (0.040) | 0.285 (0.257) | 0.288 (0.265) |
| 10 | −0.057 (0.067) | −0.001 (0.076) | 0.375 (0.362) | 0.268 (0.422) |
| 20 | −0.098 (0.070) | −0.071 (0.102) | −0.013 (0.364) | 0.403 (0.762) |

*Notes:* Estimates from Eq. (6). Dependent variable: cumulative future S&P 500 return $Y_{t,h}$. Tail-event stress: $S_t = 1$ when the date-$t$ S&P 500 return is in the lower 5% tail of the sample. Sample: daily data, December 2, 2014 to February 27, 2026, $N = 2160$. *DML-LP columns use the orthogonalised stress-interaction estimator in Eq. (7). Newey–West standard errors in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively. No coefficient in Panel B passes the Bonferroni-adjusted 1% threshold within its panel ($\alpha/5 = 0.01$), indicating that the data do not support a stable safe-haven interpretation under tail-event stress for either asset under the orthogonalised DML estimator.

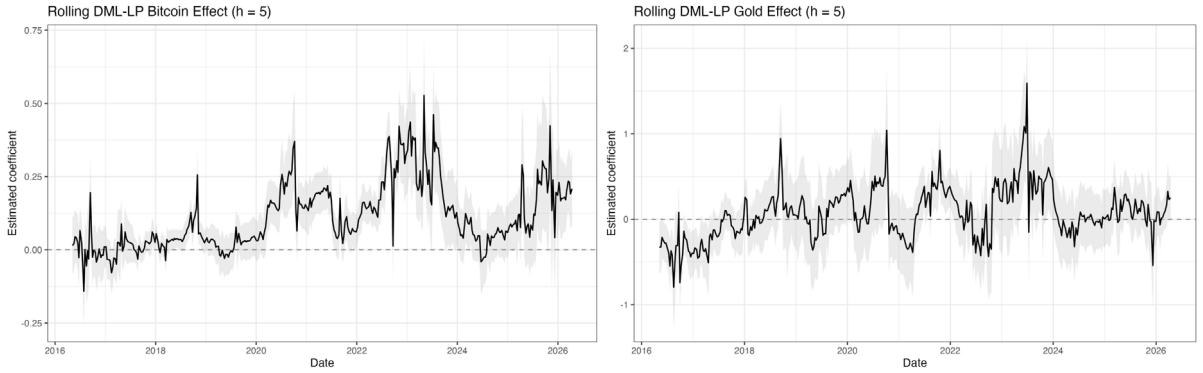
For Gold, the VIX-based stress evidence under the orthogonal DML estimator is uniformly imprecise. The stress interaction is positive at most horizons and negative at none, but no horizon clears conventional significance, let alone the Bonferroni-adjusted threshold. We do not interpret this as evidence *against* a defensive role for Gold; we interpret it as evidence that, given the data and methods, we cannot identify a stable stress-contingent hedge pattern.

The tail-stress specification in Table 5 delivers a more nuanced picture but not a more favourable one. Under the orthogonal DML estimator, none of the Bitcoin or Gold tail-stress interactions clear conventional significance, and none survive the Bonferroni adjustment. The convergence between OLS and DML in tail-stress periods is itself reassuring, since it strengthens the credibility of the headline VIX-stress finding in which they also agree directionally.

Economically, the failure of the safe-haven property for Bitcoin during VIX-defined stress aligns with a flight-to-liquidity mechanism: when equities crash, leveraged participants face margin calls, forcing liquidation of liquid alternative assets and mechanically driving procyclical comovement precisely when a safe haven is needed most.

### 5.4. Time variation

The rolling DML-LP estimates at horizon $h = 5$ (Fig. 2) demonstrate that the relation between these assets and future equity outcomes is not constant over time. Bitcoin fluctuates around zero early in the sample and becomes persistently positive during several later subperiods. Gold exhibits wider swings, including both negative and strongly positive episodes. If the underlying relation




(a) Bitcoin (b) Gold

**Fig. 2.** Rolling DML-LP estimates at horizon $h = 5$.
*Notes:* Rolling DML-LP coefficient $\theta_5$ from Eq. (5) estimated on a 252-trading-day rolling window. Horizontal axis: window end date. Vertical axis: estimated coefficient. Sample: daily data, December 2, 2014 to February 27, 2026. Both assets display substantial time variation, but the dominant pattern is one of instability rather than a persistent defensive role.

**Table 6**
Portfolio performance grid (subset)

| Allocation | Annual return | Volatility | Sharpe ratio | Sortino ratio | Max drawdown |
|---|---|---|---|---|---|
| S&P 500 only | 0.124 | 0.177 | 0.662 | 0.936 | 0.225 |
| S&P + 10% Gold | 0.125 | 0.161 | 0.734 | 1.043 | 0.207 |
| S&P + 10% Bitcoin | 0.136 | 0.185 | 0.689 | 0.958 | 0.251 |
| S&P + 10% Gold + 10% BTC | 0.137 | 0.171 | 0.753 | 1.046 | 0.234 |

*Notes:* Daily returns aggregated to annual statistics using 252 trading days. Sample: daily data, December 2, 2014 to February 27, 2026. The full grid spans Gold weights $w_G$ and Bitcoin weights $w_B$ over $\{0, 0.025, 0.05, 0.10, 0.15, 0.20\}$ subject to $w_G + w_B \leq 0.50$, with the residual weight on the S&P 500. The Sharpe ratio is computed using a zero risk-free rate; the Sortino ratio uses downside deviation as the denominator; maximum drawdown is the largest peak-to-trough decline in cumulative portfolio value over the sample. Fig. 3 visualises the maximum-drawdown surface across the full grid. Authors' calculations from Yahoo Finance and FRED data.

is time-varying, then different sample windows can generate different conclusions even with similar methodology; nonetheless, the dominant pattern remains one of instability rather than a persistent defensive role.

### 5.5. Portfolio evidence: a multi-allocation grid

To complement the regression analysis, we evaluate whether the statistical relationships translate into economically meaningful hedging benefits. We report a full grid over Gold weights $w_G$ and Bitcoin weights $w_B$ across the set $\{0, 2.5\%, 5\%, 10\%, 15\%, 20\%\}$, with the residual invested in the S&P 500. For each weight pair we compute the annualised return, the annualised volatility, the Sharpe ratio, the Sortino ratio, and the maximum drawdown over the 2014–2026 sample. Table 6 reports a representative subset; the full grid is plotted in Fig. 3.

Two patterns emerge. First, Gold's Sharpe-ratio gain is monotone but modest; the marginal benefit drops sharply beyond approximately 10%, consistent with Gold acting as a capacity-limited diversifier rather than as portfolio insurance. Second, Bitcoin's drawdown cost is monotone in its weight: any allocation above roughly 5% raises the maximum drawdown above the equity-only benchmark, even when paired with a non-trivial Gold position. Read alongside the regression evidence, this picture is consistent with our central interpretation: Gold is at best a weak, capacity-limited diversifier, while Bitcoin behaves as a return-enhancing, drawdown-amplifying risk asset. These patterns align with the precious-metal shelter analysis of Bredin et al. (2017) and are more pessimistic than the COVID-period findings of Conlon and McGee (2020), Conlon et al. (2020) for cryptocurrencies, plausibly because our longer post-2020 sample contains additional joint equity–crypto drawdowns.

To examine whether the modest full-sample diversification benefit of Gold actually materialises when it is needed, Table 7 reports the same four allocations conditional on the realised stress state. The diversification benefit of Gold is real but small, and it persists in stress states: during VIX-stress days the Sharpe ratio improves from $-1.29$ for the equity-only portfolio to $-1.21$ with 10% Gold and to $-1.02$ with the combined 10% Gold + 10% Bitcoin allocation; the maximum drawdown falls from 61.9% to 56.7% and 52.8%, respectively. Nonetheless, no allocation approaches anything that could reasonably be called crash insurance: under tail stress all four portfolios lose almost their entire value within the subsample.

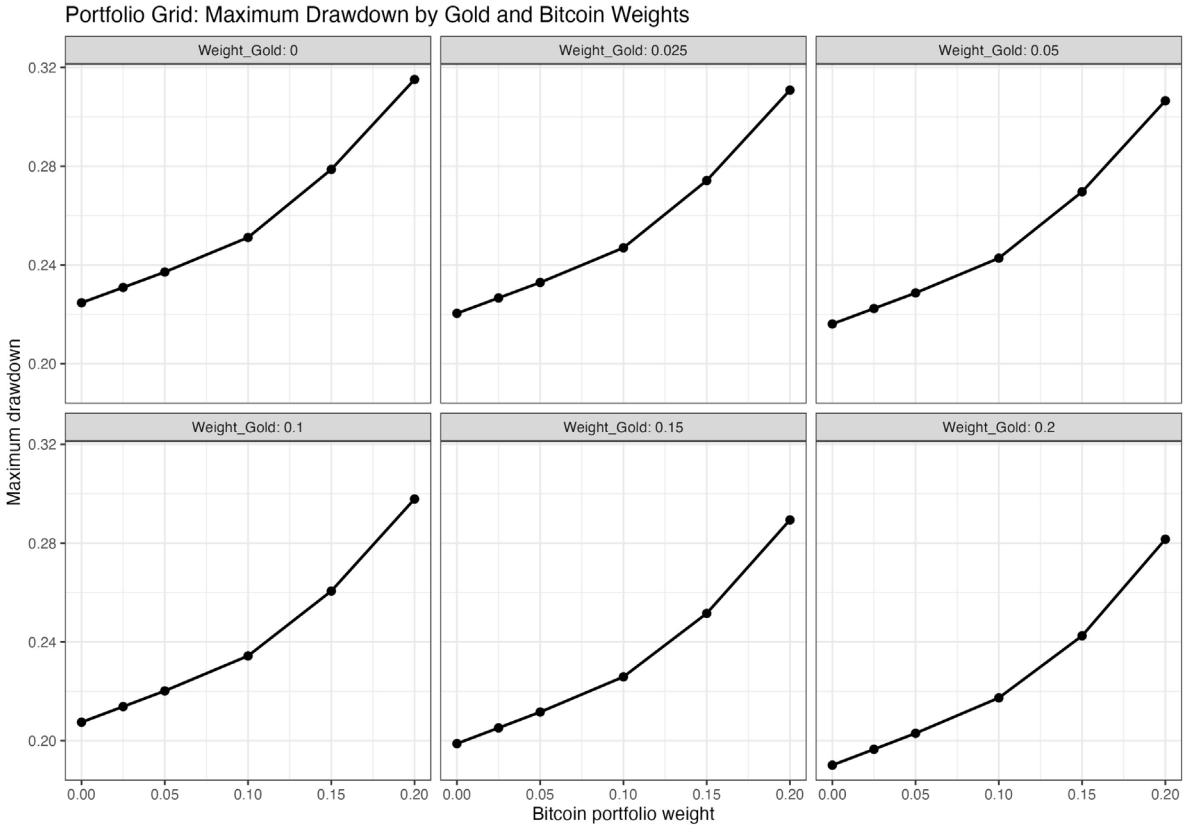



**Fig. 3.** Maximum drawdown across the $(w_G, w_B)$ portfolio grid.
*Notes:* Maximum drawdown surface across Gold weights $w_G$ and Bitcoin weights $w_B$, each over $\{0, 0.025, 0.05, 0.10, 0.15, 0.20\}$, with the residual weight on the S&P 500. Sample: daily data, December 2, 2014 to February 27, 2026. Maximum drawdown is the largest peak-to-trough decline in cumulative portfolio value over the sample.

### 5.6. Monte Carlo simulation

The Monte Carlo simulation described in Appendix A evaluates the finite-sample behaviour of our estimators under conditions relevant to the empirical application. Two key results emerge: the identification rate remains at one throughout the reported designs, but the root mean squared error rises sharply with both the forecast horizon and the degree of state-dependence (Fig. 4). The simulation is best interpreted as a result concerning finite-sample precision rather than strict non-identification, and it contextualises the wide uncertainty bands in our empirical long-horizon and stress-interaction estimates.

### 5.7. Discussion

Taken together, the results point to a clear economic conclusion: in the common sample, neither Gold nor Bitcoin behaves as a robust form of insurance against U.S. equity risk. The return-based, downside-based, and contemporaneous Baur–Lucey specifications all deliver the same qualitative answer; the DML exercise removes the alternative explanation that richer controls would have rescued the hedge interpretation, since OLS-LP and DML-LP agree closely at short horizons and DML preserves the sign at medium and long horizons. The stress-state evidence reinforces this reading. The portfolio analysis confirms the economic counterpart: adding 10% Gold cushions the VIX-stress Sharpe deterioration from $-1.29$ to $-1.21$ and the drawdown from 61.9% to 56.7%, but no allocation approaches anything that could be called crash insurance.

Two economic mechanisms are consistent with this picture. The first is liquidity demand: during distress, investors facing redemptions, leverage constraints, or margin calls liquidate the most tradable assets, generating positive comovement precisely when a safe haven should decouple. The second is the financialisation of Gold through ETFs, derivatives, and tactical asset allocation, so its short-run dynamics reflect cross-asset rebalancing as much as traditional safe-haven demand. For Bitcoin the interpretation is sharper: the evidence is more consistent with a speculative, high-beta macro-financial asset than with “digital gold”.

**Table 7**
Portfolio performance, conditional on realised stress state.

| Sample | Allocation | Annual return | Sharpe | Max drawdown |
|---|---|---|---|---|
| *Panel A. VIX-defined stress (upper quartile)* | | | | |
| Stress | S&P 500 only | −0.313 | −1.286 | 0.619 |
| Stress | S&P + 10% Gold | −0.275 | −1.211 | 0.567 |
| Stress | S&P + 10% BTC | −0.288 | −1.117 | 0.586 |
| Stress | S&P + 10% Gold + 10% BTC | −0.248 | −1.020 | 0.528 |
| Non-stress | S&P 500 only | +0.324 | +2.481 | 0.131 |
| Non-stress | S&P + 10% Gold | +0.302 | +2.550 | 0.119 |
| Non-stress | S&P + 10% BTC | +0.327 | +2.347 | 0.121 |
| Non-stress | S&P + 10% Gold + 10% BTC | +0.305 | +2.367 | 0.109 |
| *Panel B. Tail-event stress (lower 5% S&P return)* | | | | |
| Tail stress | S&P 500 only | −0.999 | −35.73 | 0.948 |
| Tail stress | S&P + 10% Gold | −0.998 | −34.22 | 0.929 |
| Tail stress | S&P + 10% BTC | −0.999 | −29.29 | 0.947 |
| Tail stress | S&P + 10% Gold + 10% BTC | −0.998 | −27.24 | 0.927 |
| Non-tail | S&P 500 only | +0.616 | +3.359 | 0.153 |
| Non-tail | S&P + 10% Gold | +0.560 | +3.406 | 0.137 |
| Non-tail | S&P + 10% BTC | +0.628 | +3.217 | 0.138 |
| Non-tail | S&P + 10% Gold + 10% BTC | +0.571 | +3.208 | 0.122 |

*Notes:* Daily returns aggregated to annual statistics using 252 trading days, computed within each subsample. Sample: daily data, December 2, 2014 to February 27, 2026. The four allocations are the same as in Table 6. VIX-defined stress: days on which the VIX is in its upper sample quartile. Tail-event stress: days on which the S&P 500 daily return is in the lower 5% tail. The Sharpe ratios in tail-stress subsamples are mechanically large in magnitude because they annualise daily statistics from a short, severe subsample; the relevant comparisons are within rather than across panels. Authors' calculations from Yahoo Finance and FRED data.

## 6. Conclusion

This paper reexamines whether Gold and Bitcoin provide hedge or safe-haven benefits against U.S. equity risk. Using daily data from 2014 to 2026, a contemporaneous Baur–Lucey-style benchmark, multi-horizon local projections under both linear and Double Machine Learning conditioning strategies, two definitions of stress, a spot-Gold robustness exercise, and a multi-allocation portfolio grid, we ask whether innovations in Gold and Bitcoin returns are associated with future S&P 500 returns and losses, and whether these relations become more defensive in periods of market stress.

The evidence provides little support for the view that either asset offers robust protection against future equity downturns. The contemporaneous and the multi-horizon designs deliver the same qualitative message; richer nonlinear conditioning does not overturn the conclusion; and the stress-state evidence is even less favourable to the safe-haven view, especially for Bitcoin under VIX-based stress. We do not claim that the data reject a defensive role for Gold, but they do not support a robust safe-haven role for Gold either. The more honest reading is that Gold acts as a weak and unstable diversifier rather than as crash insurance, and that Bitcoin behaves as a speculative macro-financial asset whose drawdowns coincide with equity stress. Economically, the results are more consistent with liquidity-driven liquidation, financialisation, and broader risk-on/risk-off portfolio dynamics than with the view of Gold or Bitcoin as reliable equity crash insurance.

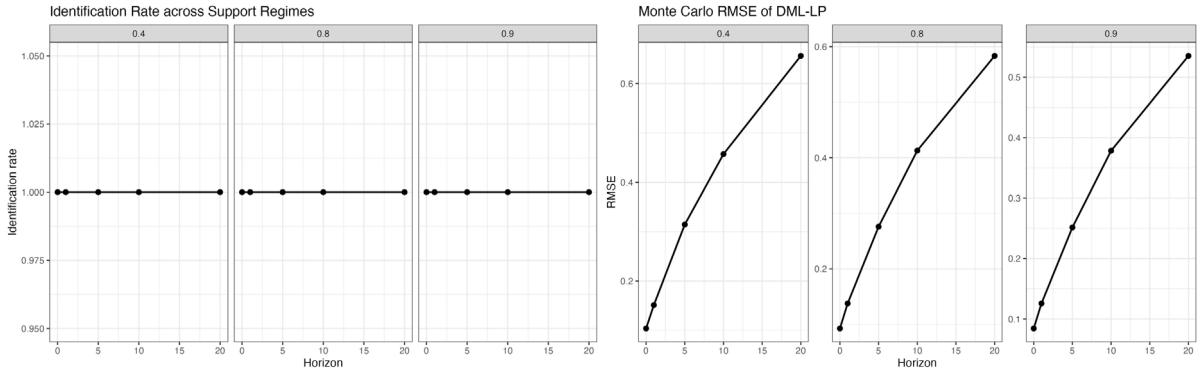
The main contribution is therefore economic rather than methodological. Our evidence sharpens the familiar conclusion in three specific ways: (i) the rejection of the safe-haven hypothesis for Bitcoin under VIX-defined stress survives a Bonferroni correction at three of five horizons, both in the canonical contemporaneous benchmark and in the orthogonalised DML stress-interaction estimator; (ii) the contemporaneous and multi-horizon designs deliver the same answer in the same sample with the same conditioning, addressing a tension that earlier studies have not had to confront; and (iii) we adopt an asymmetric position the literature rarely makes carefully: Bitcoin clearly fails the safe-haven test at multiple-testing-corrected significance, while Gold's evidence is too imprecise under the orthogonalised estimator to reject a defensive role.

These findings sit alongside a recent FRL literature on conditional safe-haven properties. Our headline rejection for Bitcoin under VIX-defined stress is consistent in spirit with Gurdgiev and Petrovskiy (2024), who argue that Gold's safe-haven role is conditional rather than universal, and with Tarchella et al. (2024) and Echaust et al. (2024), who document unstable hedge properties for cryptocurrencies. We differ from Chibane and Janson (2025), who find Bitcoin protective against geopolitical crash risk: under VIX-defined U.S.-equity stress over a longer 2014–2026 sample and with multiple-testing correction, the protection does not appear.

Future work might investigate whether the defensive properties of Gold and Bitcoin differ across monetary regimes, inflation environments, or systemic liquidity stress, and whether intra-day data deliver a sharper picture of the flight-to-liquidity dynamics suggested by our results.

## CRediT authorship contribution statement

**Marcelo J. Villena:** Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Conceptualization. **Mauricio G. Villena:** Writing – original draft, Validation, Methodology, Investigation, Formal analysis, Conceptualization.




(a) Identification rate (b) RMSE by horizon

**Fig. 4.** Monte Carlo evidence on finite-sample behaviour of the DML-LP estimator.
*Notes:* Identification rate (left) and root mean squared error (right) of the DML-LP estimator across 250 Monte Carlo draws, plotted against the forecast horizon $h$ and the state-dependence parameter $\lambda \in \{0.4, 0.8, 0.9\}$. Sample size per draw: $n = 600$. Identification remains intact across the reported designs, but estimation error rises with the forecast horizon and tighter effective support.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Appendix A. Monte Carlo simulation design and K-fold robustness

This appendix details the data-generating process underlying the Monte Carlo evidence in Section 5.6. We generate $n = 600$ observations and 250 simulation draws. The macro-financial state variables $X_{1t}, X_{2t}, X_{3t}$ are stationary AR(1) processes (e.g., $X_{it} = 0.7X_{i,t-1} + \eta_{it}$). The candidate hedge asset return $D_t$ is generated as a mixture of an idiosyncratic shock $\nu_t$ and a nonlinear function of the state $f(X_t)$:

$$D_t = (1 - \lambda_t)\,\nu_t + \lambda_t f(X_t).$$

The parameter $\lambda_t \in [0, 1]$ controls the degree of state dependence. We evaluate the estimators across a grid of $\lambda_t \in \{0.4, 0.8, 0.9\}$. The target outcome is

$$Y_t = \theta_{\text{true}} D_t + \delta_{\text{true}}(D_t \times S_t) + r(X_t) + \varepsilon_t + c_t,$$

with $\theta_{\text{true}} = 0$ and $\delta_{\text{true}} = 0$ (matching the empirical null of no hedging), $r(X_t)$ a nonlinear confounding function, $S_t$ a stress indicator triggered by adverse state realisations, and $c_t$ an occasional crash draw. Cumulative horizon outcomes $Y_{t,h}$ are constructed by summing $Y_{t+j}$ from $j = 0$ to $h$.

### *K-fold cross-fitting robustness*

To verify that our baseline DML-LP estimates are not artefacts of the specific cross-fitting partition, we re-estimate the four baseline specifications using $K \in \{2, 5, 10\}$ cross-fitting folds. Our main-text results use the chronological 50/50 split ($K = 2$); larger $K$ values use $K$ chronological folds with cross-fitted nuisance models. Table 8 reports the four sets of estimates side by side.

## Appendix B. Spot-gold robustness

This appendix reports the spot-Gold counterpart of the GLD-based results in Tables 2–4. We use the continuous COMEX Gold futures contract (Yahoo Finance ticker GC=F) as the alternative spot-Gold reference.[2] The spot series is aligned to the trading-day cross-section of the GLD-based sample.

[2] We originally intended to use the FRED LBMA daily fixing series (GOLDAMGBD228NLBM). That series was withdrawn from FRED on 31 January 2022 when the ICE Benchmark Administration data licence ended. Continuous COMEX Gold futures are the standard practitioner alternative and are highly correlated with the LBMA fixing during the period of overlap.

**Table 8**
DML-LP K-fold cross-fitting robustness.

| | Return-based | | | Downside-based | | |
|---|---|---|---|---|---|---|
| Horizon | $K = 2$ | $K = 5$ | $K = 10$ | $K = 2$ | $K = 5$ | $K = 10$ |
| *Panel A. Bitcoin* | | | | | | |
| 0 | 0.068*** (0.015) | 0.063*** (0.016) | 0.064*** (0.015) | −0.070 *** (0.014) | −0.063 *** (0.015) | −0.064 *** (0.016) |
| 1 | 0.065*** (0.009) | 0.060*** (0.009) | 0.061*** (0.009) | −0.065 *** (0.009) | −0.059 *** (0.009) | −0.062 *** (0.009) |
| 5 | 0.081*** (0.016) | 0.079*** (0.016) | 0.088*** (0.016) | −0.081 *** (0.016) | −0.078 *** (0.016) | −0.090 *** (0.016) |
| 10 | 0.056*** (0.019) | 0.061*** (0.018) | 0.073*** (0.020) | −0.060 *** (0.019) | −0.056 *** (0.018) | −0.076 *** (0.020) |
| 20 | 0.047 (0.025) | 0.063** (0.025) | 0.087*** (0.027) | −0.043 (0.027) | −0.062 ** (0.025) | −0.092 *** (0.028) |
| *Panel B. Gold* | | | | | | |
| 0 | 0.099** (0.042) | 0.072 (0.047) | 0.076 (0.048) | −0.099 ** (0.043) | −0.075 (0.047) | −0.074 (0.047) |
| 1 | 0.105** (0.050) | 0.088 (0.056) | 0.094 (0.056) | −0.109 ** (0.052) | −0.093 (0.057) | −0.095 * (0.055) |
| 5 | 0.094 (0.061) | 0.075 (0.064) | 0.088 (0.065) | −0.095 (0.061) | −0.082 (0.064) | −0.091 (0.065) |
| 10 | 0.193** (0.082) | 0.176** (0.086) | 0.191** (0.088) | −0.202 ** (0.082) | −0.183 ** (0.087) | −0.197 ** (0.088) |
| 20 | 0.222* (0.128) | 0.260** (0.123) | 0.306** (0.134) | −0.257 ** (0.123) | −0.257 ** (0.126) | −0.296 ** (0.134) |

*Notes:* DML-LP coefficient on the asset return, in the return-based ($Y_{t,h}$ = cumulative S&P 500 return) and downside-based ($Y_{t,h}$ = cumulative S&P 500 loss) specifications. Sample: daily data, December 2, 2014 to February 27, 2026, $N = 2160$. Newey–West standard errors in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively. Cross-fitting uses $K$ chronological folds; the main-text results in Tables 2 and 3 use $K = 2$. Minor differences between the $K = 2$ column here and the DML-LP estimates in Tables 2–3 reflect the stochastic nature of Random Forest estimation; setting an explicit random seed yields exact replication. At the horizons that carry the headline message ($h = 0, 1, 5$), estimates are essentially invariant to $K$; at long horizons estimates become slightly stronger as $K$ increases, consistent with the standard intuition that larger $K$ provides more training data per fold.

**Table 9**
Spot-Gold robustness (continuous COMEX Gold futures, Yahoo GC=F)

| | Return-based | | Downside-based | VIX-stress |
|---|---|---|---|---|
| Horizon | OLS-LP | DML-LP | DML-LP | DML-LP $\delta_h^*$ |
| 0 | +0.042 (0.027) | +0.060 (0.041) | −0.061 (0.041) | +0.069 |
| 1 | +0.065* (0.034) | +0.073* (0.043) | −0.075* (0.044) | +0.094 |
| 5 | +0.048 (0.044) | +0.054 (0.047) | −0.051 (0.046) | +0.109 |
| 10 | +0.063 (0.058) | +0.113* (0.060) | −0.096 (0.060) | +0.249 |
| 20 | +0.064 (0.076) | +0.139 (0.088) | −0.114 (0.086) | +0.118 |

*Notes:* Estimates using the continuous COMEX Gold futures contract (Yahoo Finance ticker GC=F) as the candidate hedge asset, in place of the GLD ETF used in the main text. Dependent variable: cumulative future S&P 500 return (return-based columns) or cumulative future S&P 500 loss (downside-based column). VIX-stress: orthogonalised DML stress-interaction coefficient $\delta_h$ from Eq. (7). Sample: daily data, December 2, 2014 to February 27, 2026, $N = 2160$. Newey–West standard errors in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively. The downside-based OLS-LP column is omitted because it is mechanically the negative of the return-based OLS-LP column under the construction $L_{t,h} = -Y_{t,h}$. None of the five VIX-stress interactions clear conventional significance and none survive Bonferroni correction. All twenty coefficients have the same sign as their GLD counterparts in Tables 2–4.

## Data availability

Data will be made available on request.

## References


Bampinas, G., Panagiotidis, T., 2017. Gold, silver, and precious metals as safe havens for stock markets. Econ. Model. 63, 48–59.
Baur, D.G., Dimpfl, T., Kuck, K., 2018. Bitcoin, gold and the u. s. Dollar—A replication and extension. Financ. Res. Lett. 25, 103–110.
Baur, D.G., Lucey, B.M., 2010. Is gold a hedge or a safe haven? An analysis of stocks, bonds and gold. Financ. Rev. 45 (2), 217–229.
Baur, D.G., McDermott, T.K., 2010. Is gold a safe haven? International evidence. J. Bank. Financ. 34 (8), 1886–1898.
Bouri, E., Molnár, P., Azzi, G., Roubaud, D., Hagfors, L.I., 2017. On the hedge and safe haven properties of bitcoin. Financ. Res. Lett. 20, 192–198.
Bredin, D., Conlon, T., Potì, V., 2017. The price of shelter—Downside risk reduction with precious metals. Int. Rev. Financ. Anal. 49, 48–58.
Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., Robins, J., 2018. Double/debiased machine learning for treatment and structural parameters. Econom. J. 21 (1), C1–C68.
Chibane, M., Janson, N., 2025. Is bitcoin the best safe haven against geopolitical risk?. Financ. Res. Lett. 74, 106543.
Conlon, T., Corbet, S., McGee, R.J., 2020. Are cryptocurrencies a safe haven for equity markets?. Res. Int. Bus. Financ. 54, 101248.
Conlon, T., McGee, R., 2020. Safe haven or risky hazard? Bitcoin during the COVID-19 bear market. Financ. Res. Lett. 35, 101607.
Dyhrberg, A.H., 2016a. Bitcoin, gold and the dollar—A GARCH volatility analysis. Financ. Res. Lett. 16, 85–92.
Dyhrberg, A.H., 2016b. Hedging capabilities of bitcoin. Is it the virtual gold?. Financ. Res. Lett. 16, 139–144.
Echaust, K., Just, M., Kliber, A., 2024. To hedge or not to hedge? cryptocurrencies, gold and oil against stock market risk. Int. Review Financial Analys. 94.
Gkillas, K., Longin, F., 2019. Is bitcoin the new digital gold? Evidence from extreme price movements. Int. Rev. Financ. Anal. 63, 73–80.
Gurdgiev, C., Petrovskiy, A., 2024. Is gold always a safe haven?. Financ. Res. Lett. 64, 105438.
He, Z., O'Connor, F., Thijssen, J., 2018. Is gold a sometime safe haven or an always hedge for equity investors?. Int. Rev. Financ. Anal. 60, 30–37.
Jordà, Ò., 2005. Estimation and inference of impulse responses by local projections. Am. Econ. Rev. 95 (1), 161–182.

Klein, T., Pham Thu, H., Walther, T., 2018. Bitcoin is not the new gold—A comparison of volatility, correlation, and portfolio performance. Int. Rev. Financ. Anal. 59, 105–116.
O'Connor, F.A., Lucey, B.M., Batten, J.A., Baur, D.G., 2015. The financial economics of gold—A survey. Int. Rev. Financ. Anal. 41, 186–205.
Reboredo, J.C., 2013. Is gold a safe haven or a hedge for the u.s. dollar? implications for risk management. J. Bank. Financ. 37 (8), 2665–2676.
Selmi, R., Mensi, W., Hammoudeh, S., Bouoiyour, J., 2018. Is bitcoin a hedge, a safe haven or a diversifier for oil price movements?. Energy Econ. 74, 787–801.
Shahzad, S.J.H., Bouri, E., Roubaud, D., Kristoufek, L., Lucey, B., 2020. Safe haven, hedge and diversification for G7 stock markets: Gold versus bitcoin. Econ. Model. 87, 212–224.
Smales, L.A., 2019. Bitcoin as a safe haven: Is it even worth considering?. Financ. Res. Lett. 30, 385–393.
Syuhada, K., Suprijanto, D., Hakim, A., 2022. Comparing gold's and bitcoin's safe-haven roles against energy commodities during the COVID-19 outbreak: A vine copula approach. Financ. Res. Lett. 46, 102471.
Tarchella, S., Khalfaoui, R., Hammoudeh, S, 2024. The safe haven, hedging, and diversification properties of oil, gold, and cryptocurrency for the G7 equity markets: evidence from the pre- and post-COVID-19 periods. Research Int. Business Finance 67 (PB).