Politecnico di Milano

Econometrics Project

Authors:
Marco Giovanni Barbero
Ettore Davide Brignoli
AgataVenturi

June 16, 2025

Contents

1 Introduction & Literature Review

In recent decades, the interrelationships among key commodities, in particular gold and oil, and equity markets have drawn growing academic interest, especially in periods of economic and geopolitical uncertainty. These relationships have been widely documented across various settings, with also some attention to their potential nonlinearity and time-varying behavior.

Our aim is to study the relationship between oil, gold, and the MSCI World Index, a proxy for global equity markets. We will adopt an incremental approach, starting with a Vector Autoregression (VAR) model, and moving toward more complex specifications, if needed, hoping to refine our results along the way. We used the studies summarized below both as an inspiration in our research and a conceptual reference, as we critically compared their results with our own during the project.

Huseynli (2023) investigates the causal relationship between gold and oil prices using daily data going from 2011 to 2022, with a specific focus on pre and post-COVID-19 periods. Through Granger causality analysis, the study finds a one-way causal link from gold prices to oil prices before the pandemic, indicating that gold may serve as a leading indicator for oil under stable conditions. However, the relationship becomes weaker after the pandemic, suggesting a disruption in traditional market linkages, due to a higher uncertainty and exogenous shocks. The study also highlights gold's traditional role as a safe-haven asset and confirms the economic relevance of these two commodities.

Arfaoui and Ben Rejeb (2017) explore the simultaneous interdependencies among oil prices, gold prices, the U.S. dollar, and international stock markets over the period 1995–2015. Using a system of simultaneous equations, the authors reveal significant direct and indirect linkages between the four markets. In particular, they find that oil prices are positively influenced by both gold and the U.S. dollar, and negatively correlated with stock market performance. Gold, is shown to be responsive to changes in oil, exchange rates, and equities. This underscoers its role as both a hedging instrument and a reactive commodity.

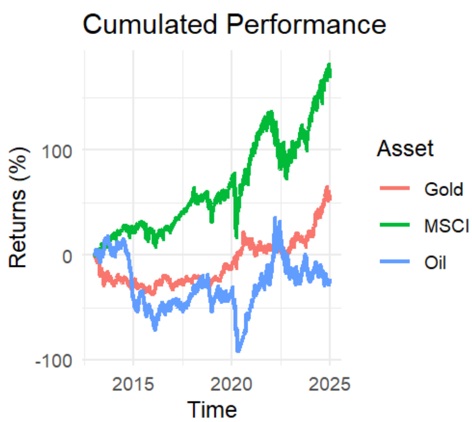
Zhao and Wang (2022) extend the literature by introducing economic policy uncertainty (EPU) and monetary policy uncertainty (MPU) as determinants of time-varying correlations among oil, gold, and stock markets in the U.S. and China. Using a DCC-GARCH t-Copula model and quantile regression, they find that U.S. policy uncertainty consistently reduces gold-stock correlations, confirming gold's function as a safe haven asset. Meanwhile, EPU and MPU amplify oil-stock correlations under certain regimes, suggesting oil's sensitivity to macro policy shifts. In particular, from Zhao and Wang (2022) we will borrow the DCC-GARCH idea which, even if not complex as their copula based, will be our most advanced model.

Karageorgou (2024), uses an ARDL model to examine the conditional relationships between gold, oil, and U.S. stock markets over sixty years. The study finds a short-run negative effect of gold on equities but no significant long-run relations, suggesting that these relationships have become weaker and become more context-dependent over time. This work underscores the value of distinguishing between short-run dynamics and long-run equilibrium relationships, an approach we also adopt with our VAR model.

Given the complexities and different findings in the literature on causality and correlations among these important asset classes, this project aims to provide a comprehensive analysis of their interactions with global markets. By using a VAR model as a baseline and exploring extensions through volatility modeling, we seek to assess how robust these relationships are across different periods and conditions.

2 Data & Methodology

Our study begins with analyzing the dataset, consisting in daily closing price observations, from the 1st of January 2013 to the 31st of December 2024, of gold spot denominated in US dollar, crude oil WTI front month futures and MSCI World Global Index.



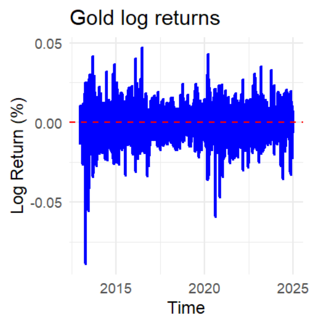
Looking at the cumulated performance of the three assets, we clearly see a huge return delivered by MSCI World Index, in great part due to a strong bull market in the US, which represents almost 70% of the index. We also see a steady increase of gold prices, particularly due to higher inflationist pressure after the COVID-19 pandemic and global landscape characterized by intense uncertainty. At last, we see an erratic path for oil prices, especially after the pandemic, reflecting a more complex macroeconomic scenery.

Some simple statistics shown below can confirm what we visually extrapolated from the graph above: oil has the largest daily fluctuations, as can be seen from its higher standard deviation. We also notice that all the three assets display a negative skewness, meaning that their distributions have long left tails; They also exhibit excess kurtosis, which implies that their return distributions have fatter tails than the normal distribution. Finally, it's worth noticing that oil has the biggest negative kurtosis, picturing a market affected by frequent supply shocks.

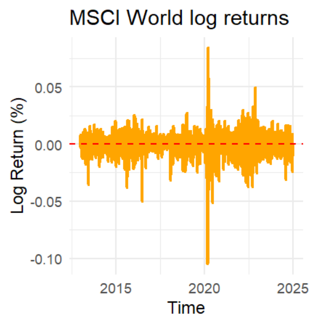
Table 1: Summary statistics of log-returns for Gold, Oil, and MSCI Index

Statistic	Gold	Oil	MSCI Index
Std. Dev.	9.31×10^{-3}	2.68×10^{-2}	9.06×10^{-3}
Skewness	-5.68×10^{-1}	-1.58	-1.09
Kurtosis	5.60	5.66×10	1.78×10

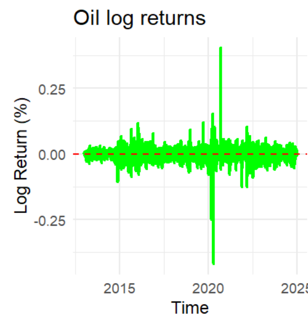
The plot below shows the log returns of each asset through time. As expected, the biggest market swings occurred during the COVID-19 pandemic, when both equity and oil crashed and quickly recovered thereafter. Gold, on the contrary, had its worst period during 2013, due to the FED "Taper" Speculation, and exhibited more limited fluctuations during the pandemic, in part confirming its status as "safe-heaven asset".



(a) Gold Log-Returns



(b) MSCI Log-Returns



(c) Oil Log-Returns

We begin our analysis with a series of preliminary tests, to evaluate both the autocorrelation within each asset (i.e., the relationship between its past and future values) and the cross-correlations between different assets.

2.1 AutoCorrelation Function (ACF Test), Partial ACF and ADF test

The autocorrelation-function test examines whether observations in a univariate time series are correlated with their own past values. For each lag k, the sample autocorrelation coefficient is estimated by comparing deviations from the series mean at time t with those at time t minus that lag.

$$\hat{\rho}_k = \frac{\displaystyle\sum_{t=k+1}^{n} (x_t - \bar{x})(x_{t-k} - \bar{x})}{\displaystyle\sum_{t=1}^{n} (x_t - \bar{x})^2} \tag{1}$$

Using the formula above, we build an hypothesis test. Under the null hypothesis we have that the autocorrelation for a fixed lag k is a white noise. We can reject the null hypothesis if the estimated autocorrelation value is large enough.

$$\begin{aligned} H_0 &: \ \rho_k = 0, \qquad \sqrt{n}\,\hat{\rho}_k \xrightarrow{d} N(0,1) \\ H_1 &: \ \rho_k \neq 0, \end{aligned} \qquad Z_k = \frac{\hat{\rho}_k}{1/\sqrt{n}} = \sqrt{n}\,\hat{\rho}_k \tag{2}$$

Reject $H_0 : |Z_k| > z_{1-\alpha/2} \iff |\hat{\rho}_k| > \frac{z_{1-\alpha/2}}{\sqrt{n}}$ For $\alpha = 0.05$: $|\hat{\rho}_k| > \frac{1.96}{\sqrt{n}}$

Unfortunately the ACF doesn't give a clear picture about the correlation of a single lag K, since it doesn't exclude the influence of all the lags k < K. So, we applied the Partial AutoCorrelation Function, where the test quantifies the direct linear relationship between a time series and its own past values, while excluding the influence of all shorter intervening lags. For each lag k, one first fits an autoregressive model of order k − 1 to the series and to its k-lagged copy, obtaining two sets of one-step-ahead residuals. The partial autocorrelation at lag k is then the sample correlation between those two residual series, which isolates the incremental effect of the k-lagged value once the effects of lags 1 through k − 1 have been removed.

Finally, the Augmented Dickey–Fuller test is based on estimating, by ordinary least squares, a dynamic regression in which the first difference of the series is regressed on its own lagged level and a finite number of lagged differences chosen to absorb serial correlation in the residuals. Under the null hypothesis of a unit root, the coefficient on the lagged level term equals zero and the series follows a nonstationary process; under the alternative, it is strictly negative and the series is stationary around any included drift or deterministic trend.

2.2 VAR

A vector autoregression (VAR) model treats a finite collection of jointly observed time-series as a finite linear function of its own and other's variable past observations. Stability requires that the feedback embodied in the lag coefficients is weak enough for all covariance moments to exist; in that case, ordinary least squares delivers consistent estimates, because the same regressors appear in every equation. In our case y_t is the vector of Gold, Oil and MSCI Index log returns at time t, A_j is the j-th row of A which is the matrix of the estimated coefficients, p is the lag of the model. It is an "hyper-parameter" and will be estimated using an AIC criteria.

$$\begin{aligned} \mathbf{y}_t &= \sum_{j=1}^{p} A_j\,\mathbf{y}_{t-j} + \mathbf{u}_t, && \mathbf{u}_t \sim \mathcal{N}(\mathbf{0}, \Sigma) \\ \Phi(z) &:= I_k - A_1 z - \cdots - A_p z^p, && \text{stationary iff } \det\big[\Phi(z)\big] \neq 0 \ \ \forall |z| \leq 1 \end{aligned} \tag{3}$$

The Grenger causality test tests whether the inclusion of past values of one variable improves the forecast accuracy of another, by lowering the mean-square forecast error. In the context of a VAR, this corresponds to testing the joint hypothesis that all lagged coefficients of the candidate explanatory variable are zero in the target equation. To conduct the test, one first estimates the unrestricted VAR system, then re-estimates it with the relevant coefficients constrained to zero. The two models are compared using a Wald or likelihood ratio (LR) test. Under the null hypothesis, the test statistic follows an asymptotic chi-square distribution, with degrees of freedom equal to the number of restricted parameters. Rejection of the null indicates that the excluded variable contains predictive information for the target variable.

2.3 Dynamic Conditional Correlation GARCH

A univariate GARCH models the log returns r_t as the sum of μ_t its conditional mean and ε_t the mean-zero innovation whose conditional variance is σ_t^2. z_t is an i.i.d. standardized shock with $\mathbb{E}[z_t] = 0$ and $\mathrm{Var}[z_t] = 1$.

$$r_t = \mu_t + \varepsilon_t, \qquad \varepsilon_t = \sigma_t \, z_t. \tag{4}$$

Where we can also define with $p = q = 1$:

$$h_t \;=\; \varepsilon_t^2 \;-\; \sigma_t^2 \;=\; \alpha_0 \;+\; \alpha\,\varepsilon_{t-1}^2 \;+\; \beta\,h_{t-1} \tag{5}$$

Since we want to represent the relationship between our assets, it is necessary an extension of the univariate GARCH called Dynamic Conditional Correlation GARCH. In the DCC Garch framework r_t, μ_t and ε_t are vectors, in our case with dimension 3×1. In particular, ε_t is the vector of innovations whose conditional covariance matrix is

$$H_t = D_t \, R_t \, D_t. \tag{6}$$

Where $D_t = \mathrm{diag}(\sigma_{1t}, \sigma_{2t}, \sigma_{3t})$ collects the individual conditional standard deviations and R_t is the 3×3 conditional correlation matrix (positive-definite with ones on its diagonal). Each σ_{it} is described by a univariate GARCH:

$$\sigma_{it}^2 = \alpha_{i0} + \sum_{j=1}^{q_i} \alpha_{ij}\,\varepsilon_{i,t-j}^2 + \sum_{k=1}^{p_i} \beta_{ik}\,\sigma_{i,t-k}^2, \qquad i = 1, 2, 3. \tag{7}$$

For asset i, σ_{it}^2 is its conditional variance; $\alpha_{i0} > 0$ sets the long-run level; α_{ij} and β_{ik} are non-negative coefficients. It is relevant to notice that stationarity requires $\sum_j \alpha_{ij} + \sum_k \beta_{ik} < 1$. Regarding the correlation modeling we begin by defining:

$$Q_t = (1 - a - b)\,\mathbb{E}\big[z_t z_t^\top\big] \;+\; a\,z_{t-1} z_{t-1}^\top \;+\; b\,Q_{t-1}, \qquad t = 1, \ldots, T. \tag{8}$$

where Q_t is a positive-definite "quasi-covariance" matrix of the standardized shocks $z_t = D_t^{-1}\varepsilon_t$, $z_{t-1}z_{t-1}^\top$ injects yesterday's cross-movements; a measures shock impact, b measures persistence, and $a + b < 1$ guarantees mean-reversion toward $\bar{Q}$. Finally, we obtain a proper correlation matrix while preserving positive definiteness, by solving:

$$R_t = \big(\mathrm{diag}\, Q_t\big)^{-1/2}\, Q_t \,\big(\mathrm{diag}\, Q_t\big)^{-1/2}. \tag{9}$$

In the case where ε_t is modeled using a multivariate Guassian, starting from the conditional law conditioned on $\mathcal{F}_{t-1}$, we have:

$$f\big(\varepsilon_t \mid \mathcal{F}_{t-1}\big) \;=\; \frac{1}{(2\pi)^{N/2}\,(\det H_t)^{1/2}}\,\exp\!\Big(-\tfrac{1}{2}\,\varepsilon_t^\top H_t^{-1}\varepsilon_t\Big). \tag{10}$$

For optimization purpose we can drop the $\frac{1}{(2\pi)^{N/2}}$; taking the log and summing over t we get:

$$\begin{aligned}\ell(\theta) &= -\frac{1}{2}\sum_{t=1}^{T}\Big[\log\big((\det D_t)^2\big) + \log\det R_t + z_t^\top R_t^{-1} z_t\Big] \\ &= -\frac{1}{2}\sum_{t=1}^{T}\underbrace{\Big[\log\big((\det D_t)^2\big) + z_t^\top z_t\Big]}_{\text{variance part}} \;-\; \frac{1}{2}\sum_{t=1}^{T}\underbrace{\Big[\log\det R_t + z_t^\top R_t^{-1} z_t - z_t^\top z_t\Big]}_{\text{correlation part}}\end{aligned} \tag{11}$$

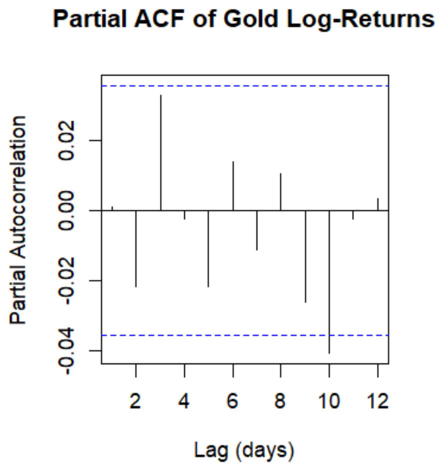
Where the last equation is obtained by summing and subtracting $z_t^\top z_t$, doing this allow us to optimize first the first term, which accounts for the ”variance part” and only after to optimize the second term which accounts for the correlation. This procedure is called QMLE, i.e. quasi maximum likelihood estimator, and was firstly introduced in this framework by Engle's (2002).

3 Results

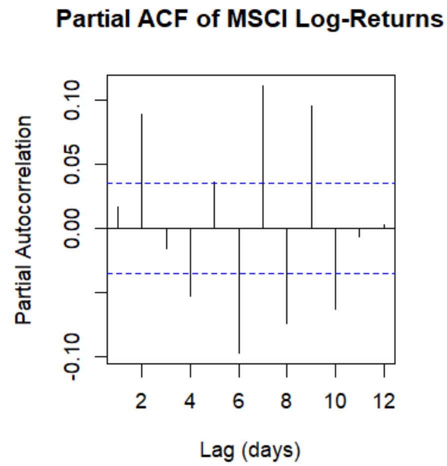
We start by looking at the autocorrelation of each asset using the partial ACF test; for simplicity the resulting coefficients are displayed in a bar chart, and approximate 95 % confidence bands are drawn around zero based on the assumption of white noise. If a bar exceeds its confidence band, the null hypothesis of no autocorrelation at that lag is rejected.

Looking at the plot below we don't have very promising results for Gold (Oil omitted since very similar). About equity we have a totally different picture since there are different significant lags. This indicates something well known to market practitioners: the importance of a trending market, which could be taken advantage of by momentum strategies.

The test statistic is the t–ratio of the estimated coefficient on the lagged level term, and its distribution under the null is nonstandard (the Dickey–Fuller distribution). So critical values specific to the inclusion of drift and trend and to the number of augmenting lags must be used. A sufficiently large negative value of this t–statistic (more extreme than the 1 percent or 5 percent critical threshold) leads one to reject the unit-root hypothesis.



(a) Gold



(b) MSCI

Figure 2: Partial autocorrelation functions for Gold and MSCI log-returns.

	Gold	Oil	MSCI
DF Statistic	-14.688	-13.375	-14.890

3.1 VAR pre-Covid

We begin our analysis by fitting a VAR model, since it is simple but effective in capturing interrelationships between variables. Since by using AIC and HQ criteria to choose the number of lags we obtained very different results on the two periods, we decided to use the criteria on the whole dataset and then perform two models with the same number of lags. So we fitted the VAR using a number of lags $p = 10$. Below we report only the cross-equation coefficients that remain significant at $p < 0.05$ when heteroskedasticity-robust standard errors are used, an hypothesis that we will test later.

Cause $\to$ Effect	Lag (days)	Coefficient	p
Oil $\to$ MSCI	1	0.018	0.036
MSCI $\to$ Gold	5	0.070	0.044
Gold $\to$ Oil	4	-0.136	0.012

The estimated VAR indicates a modest but statistically reliable sequence of lead-lag effects among the three markets. The daily increase in crude oil prices tends to be mirrored almost immediately by global equity moves, implying that energy functions as a near-term indicator macro conditions.

Whereas from equities to gold: when stocks perform strongly, the model predicts a delayed uptick in the metal's returns roughly one week later. This could be interpreted as investor's reallocation of recent gains into a perceived more stable asset like gold. Finally a surge in gold is followed, after few days, by a mild pull-back in oil prices, a pattern also identified by Huseynli (2023).

We also performed a Granger test in order to assess the explanatory power in a more comprehensive way.

Cause → Effect	Statistic	p
Gold → Oil, MSCI	1.1156	0.3243
Oil → Gold, MSCI	0.9969	0.4620
MSCI → Gold, Oil	1.0943	0.3471

The tests show that, in the pre-COVID subsample, none of the three assets adds incremental forecasting power for the others once their own ten-day histories are taken into account. In fact, the F-statistics for Gold, Oil, and MSCI are all close to 1 and their p-values are well above conventional thresholds, so the null hypotheses of "no Granger causality" cannot be rejected. So any individual cross-lag coefficient that happened to look significant in the VAR is not strong enough, when considered together with the other lags.

3.2 VAR post-Covid

Regarding the post-covid dataset we fitted two VAR models with different lags. The first with $p = 1$ minimizes the HQ criteria and the obtained results are similar to the pre-covid period, with strong explanatory power coming from Oil. Whereas the fitted VAR(10) minimises the Akaike (AIC) and gives a flipped picture about the previous role of Oil with instead using Gold as relevant leading indicator. It is curious to notice that the explanatory power of Gold over Oil was also founded by Huseynli (2023) in the pre-Covid period.
Both systems are covariance–stationary (largest companion-matrix root < 0.90), but univariate ARCH-LM tests on the raw returns reject homoskedastic innovations at $p < 10^{-15}$; all inference below therefore relies on heteroskedasticity-robust (White) covariance matrices.

Model	Cause → Effect	Lag (days)	Coefficient	p
VAR(3)	MSCI → Oil	1	−0.259	0.002
	Oil → MSCI	3	0.030	0.003
VAR(10)	MSCI → Oil	6	0.329	< 0.001
	Gold → Oil	4	−0.133	0.014

Block Granger tests indicate that MSCI leads the other two markets in both specifications. Oil shows significant predictive content only over the first three trading days, whereas Gold's influence emerges only when seven additional weekly lags are included. Instantaneous causality tests (not shown) strongly reject the null of diagonal residual covariance for every pair, confirming common same-day shocks.

	VAR(3)		VAR(10)	
Cause $\rightarrow$ Effect	F	p	F	p
Gold $\rightarrow$ (Oil, MSCI)	1.67	0.125	2.22	0.0014
Oil $\rightarrow$ (Gold, MSCI)	2.86	0.009	1.11	0.332
MSCI $\rightarrow$ (Gold, Oil)	3.62	0.0014	3.37	5.7×10^{-7}

Impulse–response analysis corroborates these findings: in VAR(3) a one-standard-deviation Oil shock raises MSCI returns for two days, while Gold responds mainly to MSCI shocks. The extra Gold$\rightarrow$Oil path in VAR(10) is economically small (peak effect below 0.07 percentage points) and disappears when lag 4–10 coefficients are removed.

In summary, the VAR(3) provides the most parsimonious and robust representation of post-COVID cross-market dynamics: it preserves the key predictive channels (MSCI and short-run Oil leadership), controls parameter proliferation, and is straightforward to extend to a VAR-GARCH framework that accommodates the pronounced conditional heteroskedasticity present in the data.

3.3 DCC-GARCH Pre-Covid

In order to account for the time-varying correlation between assets a DCC-GARCH is a reasonable choice to improve our results. Since it is well known that negative return shocks typically raise future volatility more than positive shocks of the same magnitude, we decided to use an Exponential GARCH as underlying driver, where we model the log of h_t

$$\log h_t = \omega + \alpha \left(\frac{|\varepsilon_{t-1}|}{\sqrt{h_{t-1}}} - \mathbb{E}\left[\frac{|\varepsilon_{t-1}|}{\sqrt{h_{t-1}}} \right] \right) + \gamma \, \frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}} + \beta \, \log h_{t-1},$$

We also decided to model the innovators z_t using a t-Student distribution in order to account for the fat tails of our assets. These last two remarks differ from the one discussed above, in which we assumed a DCC-GARCH with multivariate Guassian distribution. The reason is that the derivation was cleaner and the underlying ideas are unchanged.

Unlike the VAR, the DCC–eGARCH model reserves all cross–market interaction for the risk block. Each asset's volatility is driven almost entirely by its own past—$\beta \approx 0.99$ for gold and oil and $\beta \approx 0.95$ for equities. A surprising result is γ being positive and highly significant. This means, in the pre-COVID sample, that positive shocks, not negative ones, inflate risk. The dynamic-correlation recursion shows the "news–impact" coefficient $a \simeq 0.018$ implies that a fresh co-movement shock lifts every pairwise correlation by roughly two percentage points, and the persistence coefficient $b \simeq 0.939$ allows almost that entire rise to linger for more than ten trading days. Thus the VAR indicates that directional predictability is weak and staggered, whereas the DCC layer shows that markets' tendency to move together jumps almost instantly after any shock and then decays only very slowly. For this reason, it is important to not rely on the VAR links for day-to-day timing, but treat the DCC parameters as crucial for value-at-risk and stress-testing: diversification evaporates quickly and returns only gradually.

Statistic	Interpretation	Estimate	p
a	Immediate correlation shock	0.018	0.014
b	Correlation persistence	0.939	< 0.001
ν_{joint}	Multivariate t tail index	7.24	< 0.001

3.4 DCC-GARCH Post-Covid

Post-COVID the same framework paints a different picture. Gold's volatility now reacts not only to the sign of the last shock but also to its absolute size $\alpha > 0$ and marginally significant, while its persistence falls to $\beta \approx 0.96$, meaning that volatility reverts more quickly. For equities, persistence actually rises $\beta \approx 0.98$ and γ is no longer statistically distinguishable from zero, suggesting that the pandemic bull market muted the asymmetry that had previously characterised MSCI risk. Oil retains its negative α and positive γ, but both effects intensify.

However, most of the change lies in the correlation dynamics. The impact parameter more than doubles $a \approx 0.042$, a multi-asset shock now adds about four percentage points to every contemporaneous correlation, twice the pre-COVID impulse, while the persistence parameter slips only slightly to $b \approx 0.928$. Correlation shocks therefore decay a little faster (half-life $\simeq$ 9 trading days instead of 11) but start from a much higher peak. At the same time the degrees-of-freedom of the multivariate Student-t copula fall from 7.24 to 6.11, meaning joint tails have grown fatter: extreme co-movements are now more likely than before the pandemic.

Statistic	Interpretation	Estimate	p
a	Immediate correlation shock	0.042	< 0.001
b	Correlation persistence	0.928	< 0.001
ν_{joint}	Multivariate t tail index	6.11	< 0.001

4 Conclusion

Our findings echo the established literature in showing that mean–level linkages among gold, oil and global equities are weak, whereas their risk co-movements are stronger and regime-dependent. Consistent with Huseynli (2021), who reports a fragile gold $\rightarrow$ oil influence before the pandemic, we likewise see gold's predictive role vanish once common volatility shocks dominate. The real post-2020 novelty lies in the risk architecture: the DCC-eGARCH estimates imply that correlation shocks now strike roughly twice as hard and decay only slightly faster, while joint tails thicken, in line with evidence that macro-policy uncertainty amplifies oil–stock co-movement Antonakakis(2022). On the univariate side, gold's volatility mean-reverts faster, equity volatility persists longer, and the pre-COVID leverage asymmetry in equities largely disappears, suggesting markets now penalise upside exuberance as much as downside fear.

5 References

[1] Huseynli, N. (2023). *Analyzing the Relationship between Oil Prices and Gold Prices before and after COVID-19.* International Journal of Energy Economics and Policy, 13(2), 373–378.

[2] Arfaoui, M., & Ben Rejeb, A. (2017). *Oil, gold, US dollar and stock market interdependencies: A global analytical insight.* European Journal of Management and Business Economics, 26(3), 278–293

[3] Athina, K. (2024). *The Conditional Relationships between Gold and Stock Market Prices: Evidence from USA* (Dissertation)

[4] Coronado, S., Jiménez-Rodríguez, R., & Rojas, O. (2018). *An Empirical Analysis of the Relationships between Crude Oil, Gold and Stock Markets.* The Energy Journal, 39, 193–208. https://www.jstor.org/stable/26606269

[5] Zhao, W.-Y., & Wang, Y.-D. (2022). *On the time-varying correlations between oil-, gold-, and stock markets: The heterogeneous roles of policy uncertainty in the US and China.* Petroleum Science. Advance online publication.

[6] Choudhry, A., Hassan, S. S., & Shabi, S. (2015). *Relationship between gold and stock markets during the global financial crisis: Evidence from nonlinear causality tests.* International Review of Financial Analysis, 41, 247–256.