

The Price of Ethics: Liquidity and Safe Haven Properties of Responsible Physical Gold ETPs

Borja Amor-Tapia

borja.amor@unileon.es

Despacho 161, Facultad de Ciencias Económicas y Empresariales. Campus de
Vegazana, 24071
Universidad de León, Spain

Paula Castro

paula.castro@unileon.es

Universidad de León, Spain

María T. Tascón

m.tascon@unileon.es

Universidad de León, Spain

Abstract

This paper examines the liquidity-resilience trade-off of Responsible Physical Gold ETPs compared to conventional physical gold investment vehicles. While gold is traditionally viewed as a uniform safe-haven asset, we identify a systematic “Price of Ethics” characterized by significantly higher bid-ask spreads for ethical vehicles. Using High-Dimensional Fixed Effects (HDFFE) and Propensity Score Matching (PSM), the findings suggest that this responsibility premium is associated with differentiated resilience during systemic shocks. While conventional gold exhibits heightened sensitivity to bond market volatility (MOVE), Responsible Gold exhibits characteristics of a buffer, showing a positive net return response and a simultaneous surge in trading volume. Our analysis further reveals a maturation shift: the liquidity of Responsible Gold ETPs reflects a stable structural premium as the market moves away from transitory attention shocks, while the association with thematic ESG attention reverses, transitioning from initially widening spreads to compressing them as ethical mandates become embedded in market dynamics. Robustness checks using the Amihud ratio confirm that while Responsible Gold commands a transaction premium, it maintains market depth. Overall, our results suggest that Responsible Gold ETPs function as specialized defensive instruments, where higher transaction costs are balanced by enhanced price stability during systemic shocks.

Keywords: Responsible Gold; Gold ETPs; Liquidity Risk; Safe Haven Assets; Bond Market Volatility; Market Maturation; Investor Attention; ESG Premium.

JEL Classification: G11, G12, G14, M14, Q02, Q56.

1 Introduction

Gold has long been recognized as a central safe-haven asset within global financial markets. Although widely regarded as both a hedge and a diversifier, the strength of these properties varies across market conditions and asset classes. A substantial body of research documents gold's ability to protect portfolios during episodes of equity market stress (Baur & Lucey, 2010; Baur & McDermott, 2010). However, the efficacy of this protection appears to be crisis-specific; Echaust & Just (2022) demonstrate that while gold successfully mitigated extreme downside risk during the 2008 financial crisis, it failed to provide similar statistically significant protection during the COVID-19 pandemic. This suggests that while its hedging capacity may extend across market regimes (He et al., 2018), its role as a safe haven is not uniform. Furthermore, gold prices respond asymmetrically to global economic policy uncertainty, rising during periods of heightened risk and displaying resilience as conditions stabilize (Bilgin et al., 2018). These findings underscore gold's longstanding role as a defensive asset in times of financial instability.

Yet the primary channel through which investors access gold has evolved considerably. In contemporary portfolio management, exposure to gold increasingly occurs through Exchange-Traded Products (ETPs), which offer liquidity, transparency, and ease of access relative to physical holdings. While the literature on ETF liquidity and its influence on underlying assets is extensive (Ben-David et al., 2017, 2018; Liebi, 2020), the emerging segment of Responsible Gold ETPs remains largely unexplored. This gap is notable given the rapid integration of sustainability considerations into global financial markets and the growing segmentation of the gold industry. Our paper directly addresses the research agenda proposed by Pattnaik et al. (2023), who emphasize that gold research is a rapidly evolving domain and call for new empirical studies that explore the underlying value drivers and the diversifying properties of gold-backed instruments in the context of modern financial challenges.

As sustainable finance gains prominence, investors are no longer concerned solely with gold's hedging properties; they increasingly demand transparency regarding the metal's provenance and adherence to environmental, social, and governance (ESG) standards. While the London Bullion Market Association (LBMA) Responsible Sourcing Program has largely become the industry standard, a distinct segment of explicitly labeled Responsible Gold ETPs has emerged. These vehicles differentiate themselves not necessarily through the physical composition of the bullion (since most conventional ETPs can hold LBMA-certified bars) but by formalizing ethical mandates, offering enhanced traceability, and actively targeting sustainability-conscious investors. The broader literature on ESG integration suggests that sustainability considerations can influence asset pricing, liquidity, and investor behavior. Meta-analytic evidence shows that the relationship between ESG criteria and financial performance is predominantly non-negative (Friede et al., 2015), and recent regulatory developments have strengthened this association by improving data transparency and disclosure (Kim & Yang, 2025). From an asset-pricing

perspective, these dynamics align with equilibrium models in which sustainability preferences affect expected returns (Pástor et al., 2021). Moreover, during turbulent periods, information spillovers often propagate from implied volatility indices (such as the VIX) toward sentiment measures, reinforcing the link between market stability and sustainable investment preferences (Yang et al., 2024).

In the context of gold, however, the integration of sustainability introduces a nuanced form of product differentiation. Although the physical properties of gold bars are identical, the responsible label is associated with segmentation within an otherwise homogeneous market. This differentiation may be consistent with distinct performance dynamics between responsible and conventional gold vehicles, particularly during episodes of systemic distress. While traditional gold instruments are well-established safe havens, they may exhibit a resilience gap under certain conditions. This raises an important empirical question: do Responsible Gold ETPs provide superior protection during periods when conventional products fail to deliver comparable hedging efficiency? In this setting, the “price of ethics” may represent not merely a transaction friction but an entry cost into a specialized defensive asset class. At the same time, responsible labeling may be associated with heterogeneous liquidity patterns, reflecting differences in investor clientele and information frictions.

To our knowledge, this is the first paper to analyze Responsible Gold ETPs as a distinct physical investment vehicle, allowing us to assess whether ethical mandates are associated with higher transaction costs, differentiated resilience, or both. Using a high-dimensional fixed-effects framework (Correia, 2017) applied to a global sample of ETPs listed across nine exchanges in North America and Europe from 2020 to 2025, we investigate the liquidity-resilience trade-off that characterizes this emerging market segment. Our analysis yields three main contributions.

First, we document a responsibility premium in transaction costs. Responsible Gold ETPs are associated with significantly wider bid-ask spreads than conventional LBMA-certified, even after controlling for fund size and trading volume. However, price-impact estimates based on the Amihud (2002) illiquidity ratio reveal that this spread premium does not correspond to a structural deficit in market depth. As the segment matures, responsible vehicles maintain depth levels comparable to traditional products, suggesting that higher entry costs coexist with sufficient liquidity for institutional investors.

Second, we evaluate whether responsible gold exhibits a differentiated safe haven role. While gold traditionally hedges equity market stress, Responsible Gold ETPs appear to display greater resilience during bond market stress, proxied by shocks to the MOVE Index. In this context, we also examine whether such resilience is accompanied by increased trading activity, a pattern that would be consistent with a maintained hypothesis of investor reallocation toward responsible vehicles and a potential “flight to responsibility” mechanism. Ethical mandates may attract a more stable investor base (Bollen, 2007), characterized by longer holding horizons and

reduced propensity for mechanical sell offs during fixed income volatility, although this remains an empirical question. Accordingly, any apparent offset between higher transaction costs and resilience during sovereign debt instability is interpreted as hypothesis driven rather than as established fact. Propensity Score Matching provides evidence that these patterns are unlikely to be explained by observable fund characteristics: conventional gold exhibits greater sensitivity to volatility shocks, whereas responsible vehicles display a more resilient profile. Finally, we examine the maturation of the responsible gold market and document a structural shift in the determinants of liquidity. During the Inception Phase (2020-2022), ESG attention was associated with wider spreads, reflecting novelty effects and information frictions. In contrast, during the Maturation Phase (2023-2025), thematic attention is associated with tighter spreads and enhances liquidity. At the same time, the responsibility premium reflects a stable structural cost, as the asset class transitions from sentiment-driven attention toward consolidated market pricing. This evolution suggests that while responsible gold continues to command higher transaction costs, its market depth and resilience increasingly align with those of conventional instruments.

The remainder of the paper is organized as follows. Section 2 reviews the related literature. Section 3 describes the data and empirical methodology. Section 4 presents the results on returns, market activity, and liquidity dynamics. Section 5 provides robustness checks using the PSM framework and an alternative liquidity measure. Section 6 concludes with implications for policy and investment practice.

2 Literature Review and Hypothesis Development

2.1 The Responsibility Premium and Liquidity Frictions

The literature on ESG-mandated assets presents mixed evidence regarding market liquidity. On one hand, the loyal investor hypothesis suggests that sustainable mandates attract a stable, long-term capital base that is less sensitive to short-term performance (Bollen, 2007; Renneboog et al., 2008). However, introducing ethical criteria into a traditionally homogeneous and fungible asset class such as gold may also generate frictions. We argue that the Responsible Gold designation can lead to market fragmentation from the broader bullion market through two mechanisms: market segmentation and information costs.

Following the Investor Base Hypothesis (Merton, 1987), in settings with incomplete information, investors trade assets that satisfy their institutional mandates or recognition criteria. By imposing environmental and social provenance requirements (Oakley, 2022), Responsible Gold ETPs contribute to segmenting the gold market. This certification regime is associated with the transformation of a standardized commodity into a differentiated product, effectively narrowing the pool of potential investors to those subject to ESG constraints (Oakley, 2022). This narrower base may reduce trading activity relative to the universal LBMA benchmark. With fewer active participants, market depth can decline, increasing price impact and reducing incentives for arbitrage.

This segmentation is reinforced by microstructure considerations. In the framework of Stoll (1989), bid-ask spreads reflect order processing, inventory, and adverse selection costs. Because the investor base for responsible gold is more concentrated, market makers face higher inventory risk. Unwinding positions may require trading with a limited set of counterparties, increasing the probability of holding undesired inventory during volatile periods. Moreover, liquidity risk plays a central role in fund trading costs and performance (Ben-David et al., 2017, 2018; Bae & Kim, 2020). The Responsible designation also involves additional information verification regarding compliance with sourcing standards. If this increases information asymmetry between specialized and non-specialized traders, liquidity providers may widen spreads to compensate for adverse selection risk (Glosten & Milgrom, 1985). However, as these assets institutionalize, the potential decline in market depth may be mitigated by the stability of the mandate-driven investor base. Consistent with equilibrium models (Pástor et al., 2021), any observed responsibility premium may reflect an equilibrium transaction-cost differential rather than a persistent lack of market depth. In this setting, the “Responsible” label operates as a segmentation device. It does not necessarily reflect a difference in the physical underlying asset, but rather is theoretically linked to a specialized investment mandate that is expected to attract a distinct, ESG-constrained clientele. Consequently, the observed liquidity and price dynamics are consistent with an equilibrium outcome of this market segmentation.

These considerations motivate our first hypothesis:

Hypothesis 1 (H1): *Responsible Gold ETPs exhibit a liquidity penalty, reflected in wider relative bid-ask spreads compared to conventional and LBMA-certified gold vehicles.*

2.2 Safe-Haven Properties and Market Stress

While gold’s role as a hedge against equity market volatility is extensively documented (Baur & Lucey, 2010; Harvey et al., 2019), recent literature suggests that its effectiveness is regime-dependent and crisis-specific (Echaust & Just, 2022). Using the VIX Index as a proxy for equity tail risk, Hood and Malik (2013) and Basher and Sadorsky (2016) find that gold’s hedging ability can dissipate during prolonged periods of market tranquility. Long-term historical analysis also indicates that conventional gold often fails to provide a robust hedge against rare macroeconomic disasters or significant consumption declines (Barro & Misra, 2016). This conventional gold gap is further corroborated by Tarchella et al. (2024), who observe that while gold functions as an undisputable diversifier for G7 equity markets, it failed to serve as an effective safe haven during the COVID-19 stress period, as its hedging costs rose significantly and its defensive property became invalid during the second wave of the crisis. This poses a challenge for institutional investors seeking reliable protection during systemic shocks, particularly in the fixed-income sector. Zhu et al. (2018) demonstrate that Federal Reserve Quantitative Easing (QE) announcements significantly impact gold prices through the portfolio rebalancing channel, while noting that gold’s role as an inflation hedge is strictly contingent upon high-inflation regimes. Such dynamics underpin the high

sensitivity of gold-bond linkages to U.S. monetary policy documented by Yunus (2020)

Recent high-frequency evidence shows that gold returns react asymmetrically to these monetary shocks. Specifically, gold exhibits a more pronounced rally following unexpectedly loose policy than a decline after tightening (Awartani et al., 2024). These delayed adjustments in both returns and volatility suggest that short-term inefficiencies persist in how the market absorbs interest rate information. Such dynamics contradict simple versions of the efficient market hypothesis and highlight why gold's role as a cross-asset stabilizer is increasingly complex. This complexity is further underscored by Li et al. (2023), whose findings reveal that geopolitical instability acts as a primary catalyst for gold volatility by eroding confidence in fiat currencies. Supporting this view, Younis et al. (2025) find that gold's integration with other markets is inherently crisis-specific, acting as a primary destination for risk flows during the recent Russia-Ukraine and Israel-Palestine conflicts.

We argue that Responsible Gold ETPs may partially mitigate this gap by attracting a specialized, mandate-driven investor base. These participants often view ethical gold as a strategic asset (de Pessemier et al., 2026; Michaud et al., 2006) and a clean alternative to duration risk. Unlike the high-turnover investors in standard ETPs, whose mechanical arbitrage can propagate liquidity shocks (Ben-David et al., 2017; Bae & Kim, 2020), Responsible Gold vehicles may attract a more stable investor base. These mandate-driven participants, subject to non-pecuniary ESG preferences, likely prioritize downside protection over immediate execution costs (Bollen, 2007; Renneboog et al., 2008). This behavior is consistent with the equilibrium framework of Pástor et al. (2021), where sustainable investors derive utility from their mandates, accepting structural cost differentials to satisfy ethical constraints.

This leads us to expect differentiated behavior during episodes of market distress. While conventional gold may exhibit the vulnerabilities identified by Barro & Misra (2016) and the safe-haven limitations documented by Echaust & Just (2022), a potential flight-to-responsibility toward ethical vehicles may be associated with greater price resilience and higher trading intensity. In this context, Responsible Gold may act as a defensive instrument during periods of geopolitical or financial uncertainty, since these assets appear sensitive to shifts in the threat environment (Ding et al., 2022; Parnes & Parnes, 2025). This pattern would be consistent with portfolio rebalancing, where investors trade off the liquidity penalty identified in H1 for enhanced tail-risk mitigation. This leads to our second hypothesis:

Hypothesis 2 (H2): Responsible Gold ETPs exhibit relatively greater safe-haven properties and attract higher trading activity during periods of market distress compared to traditional gold vehicles.

2.3 Market Maturation and the Shift in Information Drivers

Financial innovation typically follows a learning-by-trading trajectory, where the informational efficiency of a new asset class evolves alongside its participant base (Lo, 2004). In the nascent stages of Responsible Gold, liquidity and price discovery are frequently dominated by limited investor attention and retail sentiment. Following Barber and Odean (2008) and Da et al. (2011), individual investors act as net buyers of attention-grabbing assets, creating a regime where search volume drives trading density. In this inception phase, the responsibility label functions as a novelty signal, attracting momentum-driven flows characterized by the naïve extrapolation bias identified in the broader ETF market (Clifford et al., 2014). As the segment matures, these attention-driven anomalies gradually dissipate through institutional arbitrage, transitioning the market toward a regime where ethical mandates are integrated into fundamental pricing dynamics rather than transitory sentiment shocks.

The interconnectedness between sentiment and volatility is acute during these turbulent periods. Information spillovers typically flow from implied volatility indices, such as the VIX, toward sentiment measures (Yang et al., 2024). Social media sentiment also influences price formation, and sentiment from experienced users possesses predictive power over market returns and cross-asset volatility (Piñeiro-Chousa et al., 2018). These signals complicate noise-trader dynamics in gold ETPs as investors rebalance portfolios between equities and safe havens based on social cues. Following Angelini et al. (2018), these reallocations are often triggered by social and political sentiment, leading to investment decisions that may be detached from economic fundamentals. As noted by Ben-David et al. (2017), the role of ETFs in price discovery is initially clouded by these sentiment-driven flows before the entry of sophisticated arbitrageurs.

As the market segment matures, the track record of Responsible Gold vehicles lengthens and the SFDR provides a standardized disclosure framework, reducing information asymmetry between participants. Following the equilibrium model of Pástor et al. (2021), once ESG preferences become embedded in investment mandates, their pricing shifts from sentiment-driven anomalies toward a persistent structural factor premium

We argue that by the Maturation Phase (2023-2025), the pricing dynamics of Responsible Gold ETPs evolve significantly. This period is characterized by notable changes in the broader drivers of gold pricing (Erb & Harvey, 2025). As the asset class matures and institutionalizes, liquidity is no longer primarily driven by idiosyncratic, sentiment-driven retail attention. Instead, the persistent costs associated with ethical sourcing and market fragmentation become structurally priced into the assets. Consequently, the initial friction-inducing influence of thematic interest recalibrates as the explicit responsibility premium consolidates. These considerations lead to our third hypothesis:

Hypothesis 3 (H3): As the market matures, Responsible Gold liquidity reflects a stable structural premium rather than transitory attention shocks.

3 Data and Methodology

3.1 Sample Selection and Data Sources

The construction of our dataset follows a multi-stage filtering process to ensure a clean identification of the responsibility premium in physically-backed gold markets. Our primary universe is sourced from the World Gold Council (Goldhub), specifically the “Gold ETF Holdings and Flows” database (December 2025). From this initial population, we isolate ETPs domiciled in North America and Europe that satisfy the requirement of physical gold backing.

To guarantee the purity of the gold liquidity signal, we implement several rigorous exclusion criteria. First, we remove all currency-hedged instruments (identified through a string-matching algorithm on fund objectives) to avoid the confounding effects of derivative overlays on bid-ask spreads. Second, we exclude multi-asset vehicles (e.g., GLTR, PHPM) that invest in baskets of precious metals, as well as open-end and closed-end funds that do not follow the ETP structure. Third, instruments lacking a valid ISIN are discarded.

A significant challenge in ETP research is the presence of cross-listed versions of identical ETPs. To address this, we develop a representative instrument selection algorithm. We group ETPs by their underlying asset structure and parent identifiers and apply a hierarchical priority regime based on the settlement currency. We prioritize the U.S. Dollar (USD) as the global benchmark, followed by CAD, EUR, CHF, and GBP in descending order. In cases where a single issuer offered multiple distinct vehicles with identical investment mandates, we selected the instrument with the longest historical track record. This selection criteria ensures maximum data density analysis while further preventing the over-representation of specific issuer-level liquidity and management policies.

Our market and macroeconomic data are synthesized from four specialized sources:

1. Asset-level Market Data: Daily prices, returns, market capitalization, and trading volumes are sourced from LSEG Data & Analytics.
2. Financial Shocks: We proxy fixed-income and equity market distress via the MOVE Index and the VIX Index, respectively, with data retrieved from Datastream and FRED.
3. Geopolitical Risk (GPR): We incorporate the Caldara and Iacoviello (2022) GPR index, utilizing the latest daily series for global threats and acts.
4. Investor Attention: To capture thematic interest shocks, we construct an investor attention proxy using Google Search Volume (GSV) for the terms “Sustainable Gold” and “ESG Investing”.

The analysis in this paper focuses on the 2020-2025 period. This temporal constraint is motivated by the inception of Responsible Gold as a distinct investable asset class, with the first thematic vehicles entering the panel in early 2020.

3.2 Classification of Gold Sourcing Standards

A key contribution of this paper is the granular classification of gold ETPs based on their underlying sourcing mandates. To ensure maximum precision, we followed a two-stage process. First, we applied a heuristic string-matching algorithm on the investment “Name” and “Objective” fields in LSEG Eikon. Second, we conducted a manual line-by-line audit of the official prospectus for any instrument initially categorized as “undisclosed” or with ambiguous metadata, independently performed by all three co-authors, with any discrepancies discussed and resolved by consensus before the final classification was established.

Following this verification, instruments are categorized into two mutually exclusive regimes:

1. LBMA-Compliant (29 ETPs): Assets that reference the London Bullion Market Association (LBMA) Good Delivery standards. This group includes funds identified via the “LBMA” tag in Eikon as well as those confirmed through our manual verification.
2. Responsible Gold (5 ETPs): Assets that explicitly commit to sourcing mandates strictly more restrictive than the baseline LBMA standards. This segment is defined by formal institutional commitments to the terms “Green Gold”, “Carbon Neutral”, “Responsible Sourcing”, or “MKS PAMP” in the fund’s documentation.

This verification process allowed us to resolve all initial ambiguities. Notably, the manual audit of the prospectus led to the reclassification of one previously “undisclosed” fund as Responsible, while the remaining ones were confirmed as LBMA-compliant. By doing so, the indicator identifies the institutional “Responsible” label that serves as a segmentation device for investors, regardless of whether these mandates reflect structural differences or strategic marketing differentiation.

3.3 Empirical Strategy

Our empirical objective is to identify the association between responsible sourcing standards and investor attention on the liquidity and pricing dynamics of gold ETPs. A critical premise of our identification strategy is that the Responsible designation is a time-invariant institutional mandate established at fund inception. Because funds do not typically switch categories based on transitory market performance, the risk of reverse causality is minimized. To ensure the internal validity of our findings, we propose a three-stage econometric approach, where each stage is mapped to a specific research hypothesis. The baseline specification for liquidity determinants is defined as follows:

$$Spread_{i,t} = \alpha + \beta_1 RESP_i + \beta_2 L.Spread_{i,t} + \beta_3 L.Volume_{i,t} + \gamma Shocks_t + \gamma X_{i,t} + \eta_r + \delta_m + \epsilon_{i,t} \quad (\text{eq. 1})$$

where $Spread_{i,t}$ represents the relative bid-ask spread of asset i on day t . $RESP_i$ is a dummy variable identifying Responsible gold ETPs relative to the LBMA benchmark. $X_{i,t}$ is a vector of time-varying fund-level controls (Log Market Value, Log Asset Age, and TER). $L. Volume_{i,t}$ and $Shocks_t$ are the lagged trading volume and current financial market shocks (VIX, MOVE, GPR), respectively. η_r and δ_m are region and month fixed effects. To capture the persistent nature of liquidity frictions, we include the lagged spread ($L. Spread_{i,t}$).

To explore the informational drivers discussed in Hypothesis 3 (H3), we augment the model with standardized Google Search Volume (GSV) indices and their interactions with asset types. This allows us to test whether liquidity is associated with transitory retail attention during the market's nascent stages:

$$Spread_{i,t} = \alpha + \beta_1 RESP_i + \sum \theta(RESP_i \times Attention_t) + \beta_2 L. Spread_{i,t} + \gamma X_{i,t} + \eta_r + \delta_m + \epsilon_{i,t} \quad (\text{eq. 2})$$

To evaluate the safe-haven properties and the associated trading activity proposed in H2, we estimate Equation 3 using two distinct dependent variables ($Y_{i,t}$): daily logarithmic returns and daily logarithmic trading volume. This dual approach allows us to identify whether the responsible label is associated with resilience in asset value during stress and a simultaneous active shift in investor participation:

$$Y_{i,t} = \alpha + \beta_1 RESP_i + \sum \theta(RESP_i \times Attention_t) + \sum \varphi(RESP_i \times Shocks_t) + \gamma X_{i,t} + FE + \epsilon_{i,t} \quad (\text{eq. 3})$$

where $Shock_t$ represents exogenous financial and geopolitical shocks. $X_{i,t}$ is a vector of controls tailored to each dependent variable specification: when $Y_{i,t}$ is the asset return, the vector includes trading volume and lagged liquidity to capture market-clearing conditions; when $Y_{i,t}$ is the trading volume, the vector incorporates returns, spread, and fund-level structural characteristics (Age, TER). FE represents appropriate fixed effects combinations: asset and year-month fixed effects (ϕ_i , δ_{ym}) when $Y_{i,t}$ is the asset return, and region and year-month fixed effects (η_r , δ_{ym}) when $Y_{i,t}$ is the trading volume to ensure the identification of the time-invariant responsibility mandate.

Finally, to test Hypothesis 3 (H3), we examine the temporal stability of our estimates. We posit that the responsibility premium and its determinants are not static but evolve as the market segment matures. To identify these shifts, we employ a subsample comparison analysis, splitting our observations into two distinct developmental regimes: an *Inception Phase* (2020-2022) and a *Maturation Phase* (2023-2025). The empirical identification of market maturation is conducted by estimating Equation (2) separately for each period. The coefficient β_1 represents the "Price of Ethics" (the liquidity gap) in each era. To test Hypothesis 3, we evaluate whether β_1 is significantly lower during the Maturation period than during Inception, signaling that the Responsible Gold segment has achieved greater market depth as it established a track record.

3.4 High-Dimensional Fixed Effects

A primary empirical challenge in our setting is the presence of multiple, non-nested fixed effects within a large high-frequency panel. Standard fixed effects models face severe computational limitations when absorbing thousands of asset-day observations across overlapping regional and monthly dimensions.

To resolve this dimensionality problem, we employ the high-dimensional fixed effects (HDFFE) estimator developed by Correia (2017). This algorithm uses an iterative multi-way absorption technique uniquely suited for financial microstructures. It provides two key methodological advantages over simple demeaning approaches.

First, it offers a numerically stable method to concurrently partial out multiple levels of unobserved heterogeneity, such as asset-specific intercepts ϕ_i and year-month interactions δ_{ym} . This is designed to isolate within-asset variation rather than being confounded by global trends or regional disparities. Second, the framework accommodates multi-way clustering of standard errors. We cluster standard errors at the asset level to address serial correlation, while the HDFFE structure itself controls for correlated shocks across geographical markets. This strategy ensures consistency in large panels characterized by heterogeneous exposure to macroeconomic conditions.

We apply a staggered absorption strategy depending on the specific identification requirements of each hypothesis. For the analysis of safe-haven returns (H2), we implement Asset Fixed Effects (ϕ_i) to absorb all time-invariant fund heterogeneity, providing the most stringent test of how assets respond to exogenous shocks. However, for the analysis of the responsibility premium (H1), capital reallocation volume (H2), and market maturation (H3), we employ Region Fixed Effects (η_r). This choice is econometrically necessary: as the “Responsible” and “LBMA” designations are time-invariant at the asset level, they would be perfectly collinear with asset-specific intercepts and thus absorbed in a standard Asset FE model. The use of η_r allows us to identify these key categorical coefficients while still controlling for broad geographical heterogeneity.

3.5 Sample Composition and Periodization

We identify January 2023 as the structural breakpoint for our maturation analysis. This demarcation is theoretically grounded in three developments: (i) the stabilization of precious metal supply chains following the 2020-2022 pandemic disruptions, (ii) the implementation of Level 2 Regulatory Technical Standards of the SFDR in Europe, which standardized ESG disclosures for financial products, and (iii) the attainment of a critical track record for the responsible gold vehicles.

Empirical validation of our sample confirms the necessity of this split. As of the cut-off date, the Responsible segment exhibited a mean operational age of only 18.05 months ($SD = 14.63$), significantly lower than the 114.73 months observed for LBMA-compliant vehicles. Notably, only 33.33% of the Responsible ETPs had surpassed the 24-month seasoning period by early 2023, whereas the LBMA

segment was almost entirely composed of mature assets (88.46%). By splitting the sample at this juncture, we isolate the learning phase of the market from its subsequent institutional stabilization.

3.6 Variable Selection and Measurement

To maintain parsimony and focus on our core hypotheses, we distinguish between our primary liquidity metrics, financial shocks, and a set of fund-specific control variables.

3.6.1 Liquidity and Transaction Costs

The primary dependent variable for our analysis of transaction costs (H1 and H3) is the Relative Bid-Ask Spread. Following the standard convention for daily frequency data, we calculate this as the daily closing bid-ask spread normalized by the daily closing price:

$$Spread_{i,t} = \frac{Ask_{i,t} - Bid_{i,t}}{Price_{i,t}}$$

This metric captures the explicit cost of execution as a percentage of the asset's value. A critical refinement in our measurement process involves the treatment of non-trading days. By implementing a regional business calendar, we distinguish between “Regional Holidays” (where the market is closed) and “Idiosyncratic Illiquidity” (days where the market is active but the specific ETP has zero trading volume). We treat regional holidays as missing observations, while idiosyncratic zero-volume days are preserved as valid indicators of asset-level illiquidity, ensuring that our spread estimates are not biased by the exclusion of low-activity periods. Consequently, for days with zero trading volume, we follow the convention of maintaining the last available bid-ask spread to avoid artificial gaps in the liquidity series while acknowledging the lack of immediate execution.

3.6.2 Financial Shocks and Control Variables

To mitigate omitted variable bias and isolate the specific drivers of the responsibility premium, our empirical framework incorporates three dimensions of external shocks. First, we proxy fixed-income market uncertainty via the MOVE Index (ICE BofA), which is derived from option-implied volatility on U.S. Treasury securities and captures expected volatility in the U.S. interest-rate market. Second, equity market uncertainty is proxied by the VIX Index (CBOE), which measures the market's expectations of near-term (30-day) volatility in the S&P 500 based on option prices. Third, we account for geopolitical risk (GPR) by incorporating the GPR Threats and GPR Acts sub-indices, following the structural decomposition proposed by Caldara and Iacoviello (2022). The GPR Threats index captures news concerning potential or threatened geopolitical events, whereas the GPR Acts index captures news concerning actual geopolitical events and actions. The use of these specific sub-indices is further justified by Ding et al. (2022), who find that political risk is not a monolithic driver; instead, they demonstrate that specific dimensions (particularly external conflicts) exert the most significant influence on gold price dynamics. To capture changes rather than the levels of these external indicators, financial and geopolitical measures are operationalized as first differences ($\Delta Shock_t = Shock_t - Shock_{t-1}$).

Investor attention (a critical driver during the inception phase) is captured through Google Search Volume (GSV). Specifically, we utilize the log-transformed ($\ln(1 + GSV)$) search terms “Sustainable Gold” and “ESG Investing” to isolate thematic interest from broader market sentiment.

To facilitate the economic interpretation of our coefficients, all financial, geopolitical, and attention shock series are standardized to a zero mean and unit variance (z -score). This standardization facilitates cross-variable comparability and allows the estimated coefficients to be interpreted as the effect associated with a one-standard-deviation change in the corresponding shock measure.

$$Z(X_t) = \frac{X_t - \mu_X}{\sigma_X}$$

In addition to these exogenous shocks, our models account for a comprehensive suite of fund-specific characteristics that influence secondary market liquidity. These include market capitalization (log-transformed), daily trading intensity (log-volume, which functions as a control for liquidity persistence and a dependent variable for capital reallocation analysis), asset seasoning (time since inception), and cost structures (Total Expense Ratio). A rigorous definition of the construction, frequency, and primary sources for all variables is detailed in Table A.1 (Appendix).

4 Results and Discussion

4.1 Descriptive Statistics and Stylized Facts

Table 1, Panels A and B, provides a comprehensive overview of the dataset. Panel A presents the summary statistics, encompassing 32,369 daily observations across 34 ETPs. The average daily return for the gold ETP market is 0.065%, with a standard deviation of 1.064%, reflecting the typical volatility of the underlying commodity. The mean bid-ask spread stands at 0.15%, although the maximum value (21.19%) reflects isolated episodes of extreme illiquidity. The market presents an average Market Value of \$8,477 million (Log Market Value of 7.52) and an average daily turnover of \$118 million (Log Trading Volume of 3.69). The Average Total Expense Ratio (TER) is 0.204%, consistent with the competitive pricing of physical gold-backed vehicles. Notably, the shock variables (VIX, MOVE, and Geopolitical indices) and Investor Attention metrics are standardized (z -scores) starting in 2000 for financial shocks and 2004 for Google Search Volume (GSV) data. This long-term normalization ensures that the observed dynamics in our 2020-2025 sample are evaluated against a comprehensive historical baseline, capturing true anomalies rather than transitory fluctuations. Regarding the Investor Attention metrics, the positive means for ESG (0.834) and Gold (0.703) indicate that during our sample period, public interest in these topics was significantly above the long-term historical average, reflecting a shift toward sustainable finance and safe-haven demand.

Panel C compares the means between LBMA-compliant benchmarks and Responsible Gold ETPs. This univariate analysis reveals a persistent and significant liquidity differential between the two groups. LBMA funds maintain a tighter average bid-ask spread of 0.146%, whereas Responsible ETPs exhibit a higher average of 0.258%. The resulting difference of -0.112 percentage points is statistically significant at the 1% level ($p < 0.01$), confirming that the Responsible Gold segment is associated with higher transaction costs for investors.

Structural disparities are also highly evident in terms of scale and activity. LBMA funds are substantially larger, with an average Market Value of \$8,838.7 million (Log Market Value of 7.673) compared to \$90.9 million (Log market value of 4.073) for responsible counterparts, representing about 1% of the average benchmark fund size. This scale translates into trading activity, where LBMA funds show a significantly higher Turnover Value (\$122.8 million vs. \$0.9 million; Log of Trading Volume 3.836 vs. 0.387). Also, Daily Returns show a statistically significant difference, suggesting that "ethical premium" may influence the daily return profile during the sample period.

Table 1

Panel A: Summary Statistics - ETP variables (2020-2025)

	mean	sd	min	max	count
Daily Return (%)	0.065	1.064	-10.63	10.70	32369
Bid-Ask Spread	0.150	0.333	0.00	21.19	32369
Log Market Value	7.524	1.922	-0.08	11.94	32369
Market Value (Millions USD)	8,477	16,568	0.9	153,279	32369
Log Trading Volume	3.693	3.189	-2.30	11.18	32369
Turnover Value (Millions USD)	118	529	0.1	71,682	32369
Log Asset Age	7.945	0.893	0.69	8.95	32369
Total Expense Ratio (TER)	0.204	0.106	0.00	0.40	32369

Panel B: Summary Statistics - Shocks and Attention (2020-2025)

	Mean	Sd	Min	Max	Obs.
Financial Stress (VIX) shock	-0.003	1.128	-10.30	11.99	34,913
Bond Volatility (MOVE) shock	-0.123	1.172	-9.57	7.08	34,913
Geopolitical Threat shock	0.006	1.110	-6.40	5.23	34,913
Geopolitical Acts shock	-0.013	1.064	-4.88	3.64	34,913
Investor Attention: ESG	0.834	0.209	0.43	1.33	34,913
Investor Attention: Gold	0.703	0.907	-0.10	3.61	34,913

Panel C: Comparison of Means (LBMA vs Responsible)

	LBMA	RESPONS	Diff	p
Bid-Ask Spread	0.146	0.258	-0.112***	0.000
Daily Return (%)	0.062	0.132	-0.070*	0.019
Log Market Value	7.673	4.073	3.600***	0.000
Market Value (Millions USD)	8,838.7	90.9	8,747.8***	0.000
Log Trading Volume	3.836	0.387	3.449***	0.000
Turnover Value (Millions USD)	122.8	0.9	121.9***	0.000

Observations	31,029	1,340
--------------	--------	-------

Difference (Diff) is LBMA - Responsible. *** $p < 0.01$.

Note: All shocks and attention indices are standardized (Z-scores) performed over the maximum available historical period to provide a consistent benchmark: from 2000 to 2025 for financial and geopolitical shocks, and from 2004 to 2025 for Investor Attention (GSV) indices. Means and standard deviations reported here correspond to the 2020-2025 estimation subsample relative to these long-term baselines. The sample includes 32,369 observations across 34 ETPs (29 LBMA-compliant and 5 Responsible).

While the univariate tests in Table 1, Panel C confirm a statistically significant liquidity differential, they provide a static snapshot of the sample period. To examine the temporal evolution of this divergence, Figure 1 plots the relevant time series across the four panels.

Figure 1. Evolution of Spreads, Trading Volume, and Market Value

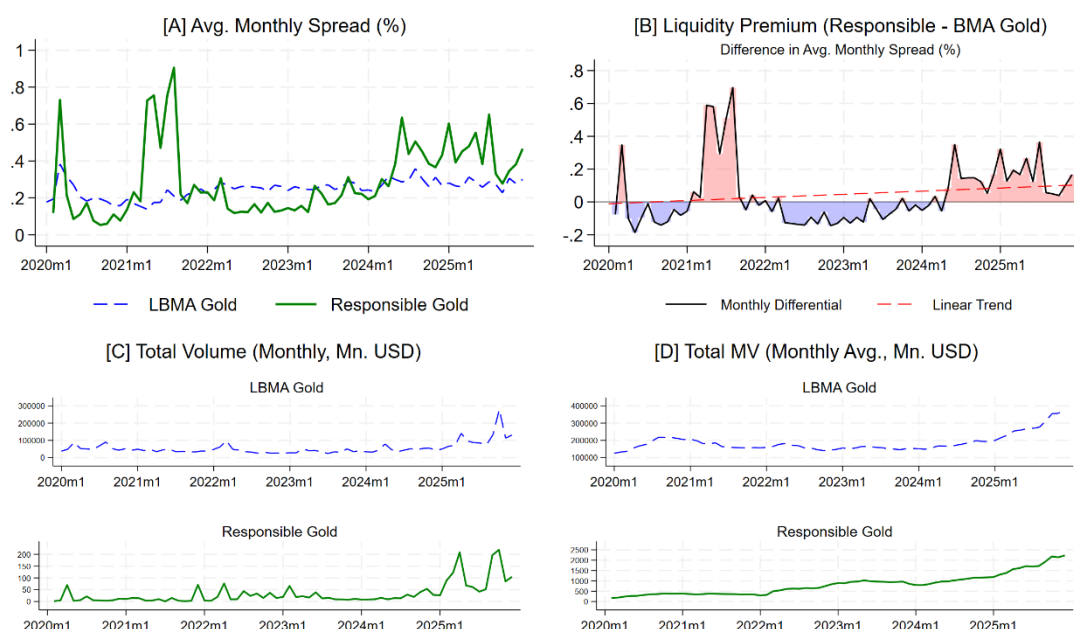


Figure 1 provides an overview of the market dynamics from 2020 to 2025. The subplots present levels, differentials, and underlying activity metrics to illustrate how market friction evolved. Collectively, the panels suggest a shift in market behavior around 2023, marking a transition from an initial consolidation phase to a period of more pronounced growth and renewed pricing pressures.

Panel [A] shows the trajectory of average monthly bid-ask spreads for both vehicle types. Standard LBMA gold instruments maintain a stable liquidity profile throughout the sample, generally remaining between 0.2% and 0.3%. In contrast, Responsible Gold ETPs display more pronounced shifts. During the 2020 to 2021 launch phase, responsible vehicles experienced significant liquidity spikes, with spreads exceeding 0.8%, which is roughly triple the baseline cost of traditional peers. Although spread levels compressed considerably between mid-2022 and

late 2023 to match the benchmark, a divergence becomes apparent from 2024 onward, with responsible spreads typically resting in the 0.4% to 0.6% range.

Panel [B] captures this friction through the time-varying liquidity premium, which is the net spread differential. The early sample is characterized by notable illiquidity, peaking near 0.7 percentage points and reflecting initial price discovery hurdles in a relatively new asset class. This was followed by a period of compression from 2022 to 2023, where the differential remained near zero or turned slightly negative, indicating temporary market alignment. However, after 2023, the premium tends to stay positive. This suggests that transaction costs for Responsible Gold have settled into a slightly higher baseline than their conventional counterparts, rather than fully converging.

Panel [C] addresses the volume dynamics underlying these spread patterns. Trading volume for LBMA-compliant ETPs operates on a much larger scale, typically between \$30,000 million and \$100,000 million monthly, with an expansion in 2025 reaching over \$200,000 million. Responsible Gold operates at a radically smaller magnitude, peaking near \$200 million per month. The volume trajectory also indicates a shift in activity. Although absolute figures remain modest, volume surges in late 2024 and 2025 point to an increase in trading interest relative to the flatter volumes seen previously.

Finally, Panel [D] shows total Market Value in USD. The LBMA segment is considerably larger, growing from approximately \$100,000 million in 2020 to nearly \$400,000 million by the end of 2025. The change in dynamics after 2023 is perhaps most apparent here. Responsible Gold accumulated capital at a very slow pace prior to 2022, but its growth rate visibly accelerates from 2023 onward. By 2023, total market value for Responsible Gold surpassed the \$1,000 million mark, and it continued climbing, rising to over \$2,000 million by the end of 2025. This inflection point suggests an evolution in investor adoption patterns and a maturation of the segment.

Whether this joint pattern of expanding adoption and persistent illiquidity represents a temporary adjustment friction or market segmentation is empirically evaluated in the econometric analysis that follows.

4.2 Multivariate Analysis: Liquidity Determinants and Investor Attention

The visual evidence of a responsibility premium in liquidity, as discussed in Figure 1, is further examined through the econometric results reported in Table 2. We employ a high-dimensional fixed-effects framework to estimate the association between investor attention and bid-ask spreads while controlling for structural asset characteristics and time-varying regional shocks.

The Baseline Model (1) indicates that the liquidity differential observed in the univariate analysis persists in a multivariate setting. After controlling for trading volume, market capitalization, expenses and asset age, Responsible Gold ETPs are associated with a statistically significant spread differential of 0.131 ($p < 0.01$) relative to the benchmark category. This pattern suggests that even after accounting for the smaller size and lower volume of these funds, the "responsible" classification itself carries a structural liquidity cost. The lagged spread term (0.369, $p < 0.01$) is positive and highly significant, indicating substantial persistence in liquidity conditions over time. Additionally, we observe that Geopolitical Acts exhibit a small but significant positive association with spreads (0.004, $p < 0.05$), suggesting that realized geopolitical events marginally increase transaction costs across the board.

Model (2) incorporates investor attention measures to test Hypothesis 1 (H1). The results point to a differentiated role of attention across categories. Attention to ESG themes is negatively associated with spreads (-0.077, $p < 0.05$), consistent with the view that thematic information demand coincides with narrower spreads at the aggregate level.

The interaction terms, however, reveal heterogeneity across gold segments. While ESG-related attention is linked to lower spreads overall (with its specific interaction with Responsible Gold, -0.11, being statistically insignificant), the interaction Responsible $\times$ Attention Gold is positive and statistically significant (0.073, $p < 0.01$). This suggest that periods of heightened general interest in gold are associated with relatively wider spreads for Responsible ETPs compared to the benchmark category. One possible interpretation is that broad gold-related attention is not accompanied by proportional liquidity improvements for this segment. This pattern may also help contextualize the localized increase in spreads observed in late 2024 and 2025 (Figure 1), where episodes of elevated gold attention coincided with a widening liquidity gap.

The Wald test for the joint significance of the interaction terms ($p < 0.001$) indicates that investor attention variables are jointly associated with cross-sectional variation in liquidity across gold ETP categories. Overall, the results suggest that thematic (ESG) attention and general gold attention are linked to liquidity in different ways, with the former associated with narrower spreads and the latter associated with relatively wider spreads for Responsible ETPs.

Table 2. Determinants of Liquidity and the Role of Investor Attention

	Model (1): Baseline	Model (2): Attention
Log Trading Volume lagged	-0.002 (0.003)	-0.003 (0.003)
Lagged Spread (t-1)	0.369*** (0.054)	0.361*** (0.053)
Responsible Gold (Dummy)	0.131*** (0.035)	0.120** (0.051)
Log Market Value	-0.001 (0.006)	-0.000 (0.006)

Log Asset Age	0.011 (0.014)	0.002 (0.014)
Total Expense Ratio (TER)	0.027 (0.062)	0.055 (0.067)
Financial Stress (VIX)	-0.000 (0.003)	-0.000 (0.003)
Bond Volatility (MOVE)	0.001 (0.002)	0.001 (0.002)
Geopolitical Threat	-0.003 (0.002)	-0.003 (0.002)
Geopolitical Acts	0.004** (0.002)	0.004** (0.002)
Attention: Gold		0.010 (0.006)
Responsible x Attention Gold		0.073*** (0.015)
Attention: ESG		-0.077** (0.031)
Responsible x Attention ESG		-0.110 (0.067)
Constant	0.014 (0.098)	0.130 (0.096)
Region Fixed Effects	Yes	Yes
Month Fixed Effects	Yes	Yes
Adj. R-squared	0.331	0.335
Observations	25917	25917
Wald test (p-val)		0.000

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: The dependent variable is the Relative Bid-Ask Spread (%). Standard errors clustered at the asset level are reported in parentheses. The benchmark category for Gold Type is LBMA-compliant. Continuous variables are standardized (Z-scores) where indicated.

4.3 Safe-Haven Analysis: Response to Geopolitical and Financial Shocks

The analysis of the safe-haven properties of gold ETPs, reported in Table 3, complements the liquidity evidence discussed in the previous section. By examining the sensitivity of daily returns to market shocks, we test Hypothesis 2 (H2), which posits that the transaction cost differential associated with Responsible Gold ETPs may be accompanied by differences in hedging performance during periods of systemic stress. The empirical specification incorporates high-dimensional fixed effects at the asset and year-month levels, allowing the estimated coefficients to capture differential return sensitivities across gold categories while accounting for time-invariant fund characteristics and common macro-financial conditions.

Model (1) examines the response to geopolitical shocks. For the benchmark (LBMA-compliant) category, Geopolitical Threats are associated with negative return effects (-0.016, $p < 0.01$), whereas Geopolitical Acts are associated with positive and statistically significant return responses (0.028, $p < 0.01$), consistent with the

conventional safe-haven interpretation of gold during realized geopolitical events. The interaction terms indicate that the return responses for Responsible Gold ETPs are not statistically significant for either threats (-0.011) or acts (-0.021). This pattern indicates that return responses for the responsible segment are statistically indistinguishable from the LBMA benchmark during geopolitical episodes, pointing to uniform safe-haven behavior across sourcing standards regarding geopolitical risk.

Model (2) evaluates the response to financial market volatility. At the aggregate level (represented by the benchmark category), gold returns are negatively associated with shocks to both equity volatility (VIX, -0.032, $p < 0.01$) and bond market volatility (MOVE, -0.041, $p < 0.01$). However, the interaction terms reveal cross-sectional differences. For Responsible Gold ETPs, the interaction with bond market volatility (MOVE) is positive and statistically significant (0.058, $p < 0.01$), suggesting more favorable return differentials relative to the LBMA benchmark during periods of elevated bond market stress. A post-estimation linear combination test (baseline MOVE coefficient plus the Responsible $\times$ MOVE interaction) indicates that Responsible Gold's total sensitivity to bond market stress is positive and marginally significant (0.017, $p = 0.089$). This finding suggests that Responsible ETPs are less exposed to bond market stress than the LBMA benchmark and that the negative response observed for the benchmark is offset in the point estimate for Responsible Gold ETPs, although the statistical evidence for the net effect is weaker than that for the interaction term. The interaction with equity volatility, by contrast, is not statistically significant. Taken together, these findings are consistent with heterogeneous hedging profiles across gold ETP segments. While the entire gold universe responds to equity stress, Responsible Gold ETPs display a more favorable return response during episodes of bond market volatility.

These results suggest that Responsible Gold ETPs combine wider spreads under normal market conditions with a comparatively more favorable response during episodes of bond market stress. Although Responsible Gold ETPs are associated with wider spreads in normal conditions, they also exhibit distinct return responses during bond market volatility episodes. This pattern may indicate differences in how gold ETP segments interact with fixed-income risk, though the underlying mechanisms cannot be directly identified within the current empirical framework.

The Wald tests for the joint significance of the interaction terms indicate that gold classification is statistically associated with differences in return sensitivities across stress dimensions. Overall, the evidence points to meaningful cross-sectional heterogeneity in safe-haven behavior within the gold ETP universe, rather than a uniform response across sourcing standards.

Table 3. Safe Haven Analysis: Asset Returns and Market Shocks

	Model (1): Geopolitical	Model (2): Financial
Geopolitical Threat (Shock)	-0.016*** (0.005)	
Responsible x Geo Threat	-0.011	

	(0.044)	
Geopolitical Acts (Shock)	0.028***	
	(0.005)	
Responsible x Geo Acts	-0.021	
	(0.028)	
Lagged Spread	0.013	0.007
	(0.019)	(0.018)
Log Market Value	0.094***	0.093***
	(0.018)	(0.018)
Log Trading Volume	-0.047***	-0.046***
	(0.008)	(0.007)
VIX (Equity Stress) Shock		-0.032***
		(0.006)
Responsible x VIX Stress		0.008
		(0.038)
MOVE (Bond Stress) Shock		-0.041***
		(0.007)
Responsible x MOVE Stress		0.058***
		(0.012)
Constant	-0.464***	-0.460***
	(0.122)	(0.125)
Asset Fixed Effects	Yes	Yes
Year-Month Fixed Effects	Yes	Yes
Adj. R-squared	0.046	0.050
Observations	26120	25928
Wald test (p-val)	0.000	0.000
Net MOVE effect, Responsible Gold (p-val)		0.017 (p=0.089)

The dependent variable is the Daily Return (%). Standard errors in parentheses. Shocks are standardized (Z-scores). LBMA-compliant is the baseline category. Wald test reports the joint significance (F-test) of the two gold type interaction terms in each model, testing whether Responsible Gold's sensitivity to the respective shocks differs from LBMA-compliant gold. Net MOVE effect reports a post-estimation linear combination test (baseline MOVE coefficient + Responsible x MOVE interaction), testing whether Responsible Gold's total sensitivity to MOVE stress is significantly different from zero. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

From an economic perspective, the magnitude of the MOVE interaction is also meaningful. Because MOVE is standardized, a one-standard-deviation increase in bond market volatility is associated with a 0.058 percentage-point increase in the daily return differential between Responsible and LBMA-compliant Gold ETPs. This more than offsets the -0.041 percentage-point response estimated for the LBMA benchmark, resulting in a net MOVE sensitivity of approximately +0.017 percentage points for Responsible Gold. For reference, the standard deviation of daily returns in the full sample is 1.064%, so the interaction effect represents approximately 5.5% of one daily return standard deviation. The statistically significant interaction indicates an economically relevant difference in the response to bond-market stress, while the positive net effect for Responsible Gold is supported at the 10% level.

4.4 Trading Activity during Market Distress

To investigate whether the outperformance of Responsible Gold ETPs during periods of market stress is supported by actual capital flows, we examine the

evolution of trading volumes. This analysis helps assess whether the observed price resilience is accompanied by greater trading activity, which would be consistent with active safe-haven reallocation rather than passive price appreciation. Table 4 presents the results of the high-dimensional fixed-effects (HDFE) regressions for daily log trading volume. We employ a region-fixed effects framework to maintain identification of time-invariant asset characteristics, facilitating a comprehensive assessment of baseline activity and heterogeneous responses to shocks.

The baseline results in Column (1) confirm that market activity is primarily driven by fund size (Log Market Value: 0.683, $p < 0.01$) and price momentum (Daily Return: -0.040, $p < 0.01$). Interestingly, the dummy for Responsible Gold carries a negative and significant baseline coefficient (-1.813, $p < 0.10$), reflecting the structural liquidity friction and the fragmented investor base discussed in Section 2.

Column (2) explores the impact of geopolitical instability. While an increase in geopolitical threats (GPR Threat) generally boosts gold trading activity (0.007, $p < 0.1$), the coefficient for the interaction term Responsible $\times$ GPR Threat (0.022) is statistically indistinguishable from zero. Similarly, the response to realized Geopolitical Acts does not show a significant differentiation for the responsible segment. This suggests that geopolitical-driven demand is absorbed uniformly across the asset class, with the responsible mandate not being associated with additional trading volume during these episodes.

The most pronounced effects emerge in Column (3) under financial shocks. We document a significant and positive interaction between the MOVE index and Responsible Gold (0.117, $p < 0.05$). In contrast, the baseline response to bond volatility for the LBMA benchmark is nearly nine times smaller (0.013, $p < 0.01$). This divergence is confirmed by the Wald test ($p = 0.008$), which rejects the null hypothesis of coefficient equality between the two gold regimes.

These findings are consistent with a possible “flight-to-responsibility” mechanism. In line with Lehnert (2022), who documents that flight-to-safety episodes are characterized by active trading and risk transfer associated with fundamental economic uncertainty rather than mere price noise, the positive volume response points to an active reallocation of capital during periods of systemic stress. During episodes of elevated bond market volatility, the results align with a pattern in which capital may be directed toward instruments with transparent sourcing mandates. However, as discussed in Section 5.6, the strength of this volume response is sensitive to the composition of the responsible segment. The absence of a comparable response to equity volatility (VIX interactions) suggests that Responsible Gold may operate as a specialized defensive instrument against fixed income and sovereign debt instability rather than as a generic hedge for equity market stress.

Table 4. Market Activity during Bond Shocks

	(1) Baseline	(2) Geopolitical	(3) Financial
Responsible Gold (Dummy)	-1.813*	-1.836*	-1.819*

	(0.958)	(0.970)	(0.962)
Log Market Value	0.683***	0.683***	0.683***
	(0.130)	(0.131)	(0.131)
Log Asset Age	0.128	0.124	0.123
	(0.262)	(0.262)	(0.263)
Daily Return (%)	-0.040***	-0.040***	-0.039***
	(0.006)	(0.008)	(0.008)
Bid-Ask Spread	-0.139	-0.128	-0.128
	(0.142)	(0.134)	(0.136)
Expense Ratio (TER)	-0.599	-0.528	-0.521
	(2.815)	(2.876)	(2.867)
GPR Threat		0.007*	
		(0.004)	
Responsible x GPR Threat		0.022	
		(0.029)	
GPR Acts		-0.003	
		(0.005)	
Responsible x GPR Acts		-0.022	
		(0.035)	
Bond Volatility (MOVE)			0.013***
			(0.005)
Responsible x MOVE			0.117**
			(0.043)
Equity Volatility (VIX)			0.006
			(0.005)
Responsible x VIX			-0.006
			(0.020)
Constant	-2.242	-2.207	-2.192
	(1.668)	(1.672)	(1.677)
Region FE	Yes	Yes	Yes
Month-Year FE	Yes	Yes	Yes
Wald Test (p-val)		0.309	0.008
Adj. R-squared	0.828	0.825	0.827
Observations	32369	26258	26049

Note: HDFE linear regressions of daily log trading volume (2020-2025).

Standard errors clustered at the fund level. Models control for size, age, returns, liquidity, and expense ratio (TER). LBMA-compliant is the baseline category.

The economic magnitude of the volume response is also substantial. Given the logarithmic specification, the Responsible × MOVE coefficient of 0.117 implies that a one-standard-deviation increase in bond-market volatility is associated with an approximately 12.4% additional increase in trading volume for Responsible Gold ETPs relative to the LBMA benchmark. By comparison, the corresponding effect for the LBMA benchmark is only about 1.3%, while the combined effect for Responsible Gold is approximately 13.9%. This magnitude supports the interpretation of the positive interaction as potentially meaningful capital reallocation rather than merely a statistically detectable difference in trading activity.

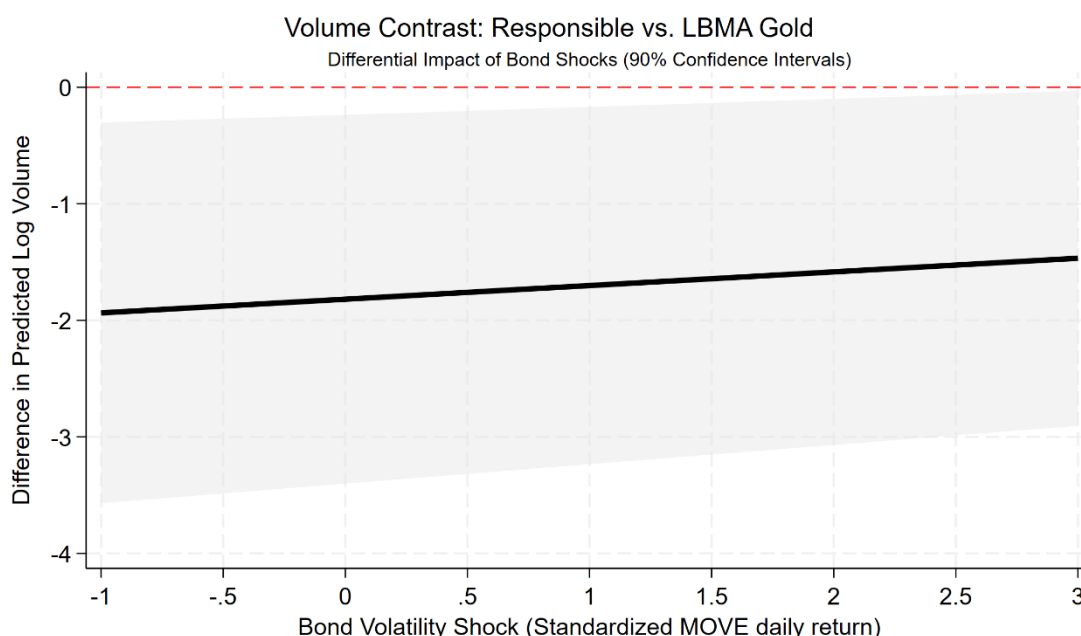
To provide a graphical illustration of the regression results, Figure 2 plots the predictive volume contrast between Responsible Gold ETPs and the LBMA-

compliant benchmark. This difference in log trading volume is calculated as a function of bond market volatility shocks. The visualization provides a descriptive depiction of the interaction effects estimated in Column (3) of Table 4.

The graphical evidence highlights a dynamic association in investor activity. The contrast remains in negative territory throughout the range. This indicates that Responsible Gold maintains a lower absolute volume compared to the LBMA segment, which is consistent with its smaller market scale. However, the slope of the differential is positive. While trading activity for the LBMA benchmark appears relatively inelastic to bond shocks, the upward trajectory of the contrast line suggests that Responsible Gold volume tends to increase more noticeably as bond volatility rises.

As bond market volatility moves from tranquil periods toward higher stress levels, the volume gap between the two segments appears to narrow. Specifically, the progression from -1 to 3 on the standardized MOVE scale shows a reduction in the differential. The 90% confidence intervals remain below the zero line, represented by the red dashed line. This confirms that the volume difference is statistically persistent across the distribution. Simultaneously, the positive slope is suggestive of a possible reallocation pattern during periods of elevated bond market risk. These patterns are consistent with a narrower volume differential between Responsible Gold and the LBMA-compliant segment during periods of higher bond market volatility.

Figure 2.



Note: Predictive volume contrast between Responsible Gold ETPs and the LBMA-compliant benchmark as a function of bond market volatility shocks. The y-axis reports the estimated difference in predicted log daily trading volume, $\ln(Vol_{Responsible}) - \ln(Vol_{LBMA})$. The x-axis reports the standardized MOVE daily return in Z-scores, with zero representing the sample mean. The solid line shows the estimated contrast and the shaded area represents 90% confidence intervals based on standard errors clustered at the fund level. Values below zero indicate lower predicted trading volume for Responsible Gold relative to the LBMA-

compliant benchmark.

4.5 Market Maturation and Structural Shifts in Liquidity Pricing

We examine these dynamics by splitting the sample into two developmental regimes: an Inception Phase (2020-2022) and a Maturation Phase (2023-2025). Table 5 presents the results of this subsample analysis.

The most notable effect is the shifting role of non-traditional liquidity drivers. During the Inception Phase, both general and thematic ESG attention influenced liquidity, though the interaction Responsible x Attention ESG exhibited a positive and highly significant coefficient (0.15, $p < 0.01$). This suggests that during the segment's early years, heightened ESG thematic interest was associated with relatively wider spreads for Responsible ETPs compared to the benchmark. However, this dynamic reverses in the Maturation Phase, where the interaction term is associated with a compression in bid-ask spreads for ethical vehicles. A similar reversal is observed in the interaction with general gold attention (Responsible x Attention Gold), which moves from -0.082 to 0.074 ($p < 0.01$). At the same time, the baseline point estimate for the Responsible Gold dummy increases from an insignificant -0.046 to a significant 0.152 ($p < 0.05$). A formal Wald test for coefficient equality explicitly rejects the null hypothesis ($F = 4.988$, $p = 0.034$), suggesting evidence of a structural shift between the two periods. From an economic perspective, this transition is consistent with a process of liquidity normalization and pricing-in of structural costs. In the inception phase, the responsibility premium was statistically indistinguishable from zero, potentially masked by the novelty of the asset class. As the segment matures, the observed premium solidifies to explicitly reflect the actual costs of ethical sourcing and market fragmentation. Rather than contradicting the maturation narrative, the combination of a structurally positive premium and its altered relationship with ESG attention suggests that Responsible Gold is moving from a sentiment-driven novelty toward a regime characterized by a stable structural premium. This is further supported by the increase in the Lagged Spread coefficient (from 0.205 to 0.425), which indicates a higher degree of liquidity persistence as these assets become more embedded in global trading regimes. Consequently, the late-sample spread differentials observed in Figure 1 reflect the crystallization of structural transaction costs rather than the idiosyncratic attention shocks that characterized the asset's early lifecycle.

Table 5. Market Maturation and Structural Shifts in Liquidity Pricing

	Model (1): Inception	Model (2): Maturation
Log Trading Volume lagged	-0.005 (0.003)	-0.004 (0.004)
Lagged Spread (t-1)	0.205*** (0.037)	0.425*** (0.088)
Responsible Gold (Dummy)	-0.046 (0.060)	0.152** (0.062)
Log Market Value	-0.005 (0.006)	0.008 (0.008)
Log Asset Age	0.016 (0.011)	-0.015 (0.029)

Total Expense Ratio (TER)	0.019 (0.073)	0.094 (0.087)
Financial Stress (VIX)	-0.002 (0.003)	0.002 (0.003)
Bond Volatility (MOVE)	0.000 (0.003)	0.002 (0.002)
Attention: Gold	-0.004 (0.018)	0.003 (0.008)
Responsible x Attention Gold	-0.082*** (0.029)	0.074*** (0.013)
Attention: ESG	-0.279*** (0.060)	-0.026* (0.015)
Responsible x Attention ESG	0.150*** (0.035)	-0.155*** (0.055)
Geopolitical Threat	-0.003 (0.002)	-0.004 (0.004)
Geopolitical Acts	0.003 (0.002)	0.005 (0.003)
Constant	0.286*** (0.077)	0.165 (0.202)
Region Fixed Effects	Yes	Yes
Month Fixed Effects	Yes	Yes
Adj. R-squared	0.295	0.365
Observations	12175	13742
Wald F-test (Diff)		4.988
Wald p-val (Diff)		0.034

Note: The dependent variable is the Relative Bid-Ask Spread (%). Standard errors in parentheses clustered at the asset level. LBMA-compliant is the baseline category. The Wald test evaluates the difference in the 'Responsible Gold' coefficient between the Maturation and Inception eras.

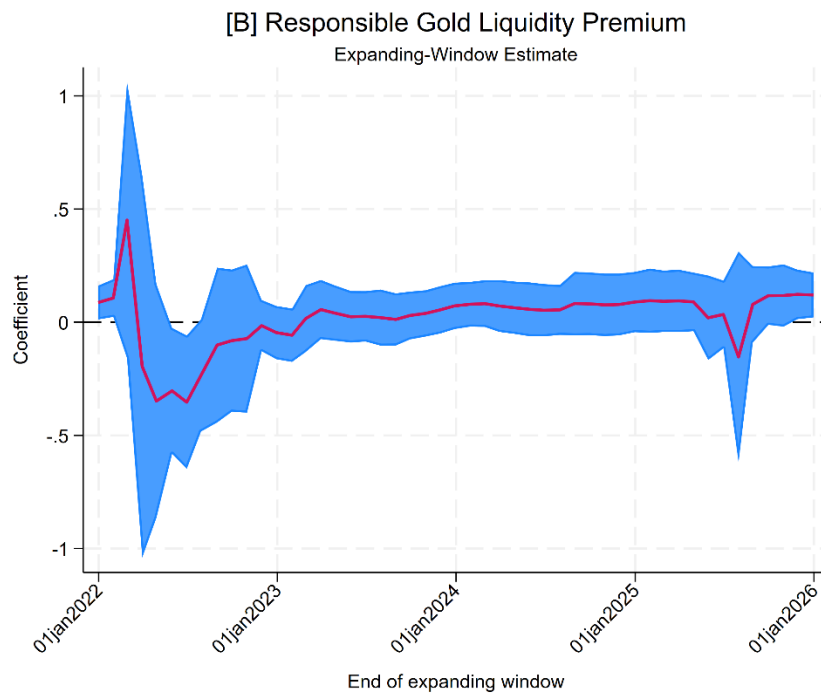
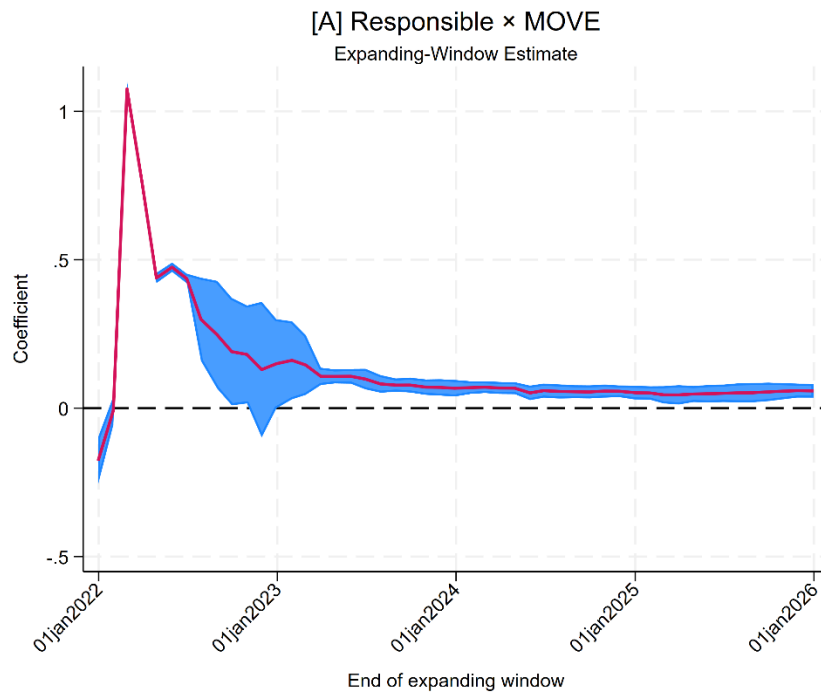
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

To complement the discrete Inception-Maturation comparison, we conduct an expanding-window analysis of two key coefficients reported in Tables 3 and 5: the Responsible $\times$ MOVE interaction in the return specification and the Responsible Gold coefficient in the liquidity model. This approach progressively incorporates new observations into the estimation and therefore provides a continuous assessment of whether the Responsible Gold liquidity premium and its sensitivity to bond-market volatility evolve gradually over time, without imposing a discrete structural break. Figure 3 reports the resulting coefficient paths.

The expanding-window estimates indicate a gradual transition rather than an abrupt change. The Responsible $\times$ MOVE coefficient in the return specification is initially highly volatile but progressively stabilizes at a positive level as the sample expands. A similar pattern emerges for the Responsible Gold liquidity premium: after substantial variation during the early observations, the coefficient converges toward a relatively stable positive level in the later sample. These patterns are consistent with the maturation hypothesis and complement the discrete differences documented in Table 5.

This evidence suggests that the liquidity-resilience trade-off is not static but recalibrates as the asset class institutionalizes. To ensure that these dynamics are not artifacts of our primary liquidity proxy, the following section provides a series of robustness checks. We employ a Propensity Score Matching framework and alternative price-impact measures to examine the stability of the documented liquidity and resilience patterns.

Figure 3. Expanding-Window Estimates of Key Responsible Gold Market Coefficients



5 Robustness Analysis

5.1 Propensity Score Matching Procedure

To address potential selection bias and endogeneity issues arising from fund-specific characteristics, we employ a Propensity Score Matching (PSM) framework. In our context, the adoption of a Responsible Gold mandate may not be random; rather, it could be correlated with structural factors such as fund size, maturity, or cost structures. If these factors also influence the ETPs' resilience during market

shocks, a simple OLS comparison would yield biased estimates. This approach allows us to disentangle the mandate-driven effect from other structural fund attributes. By comparing funds with similar size, age, and costs, we ensure that the responsibility premium and the safe-haven resilience are associated with the specific “Responsible” label rather than being an artifact of fund-level selection bias.

Following the methodology of Rosenbaum and Rubin (1983), we first estimate a logistic regression model to calculate the propensity score, i.e., the conditional probability of a fund being categorized as Responsible ($Treat = 1$) based on its observable characteristics:

$$P(X) = Pr(Treat = 1|X) = \frac{\exp(\beta X)}{1 + \exp(\beta X)}$$

Where X is a vector of pre-treatment covariates including Log Market Value (as a proxy for liquidity and scale), Log Asset Age (to account for fund maturity), and the Total Expense Ratio (TER). We also include regional fixed effects to control for geographical market variations.

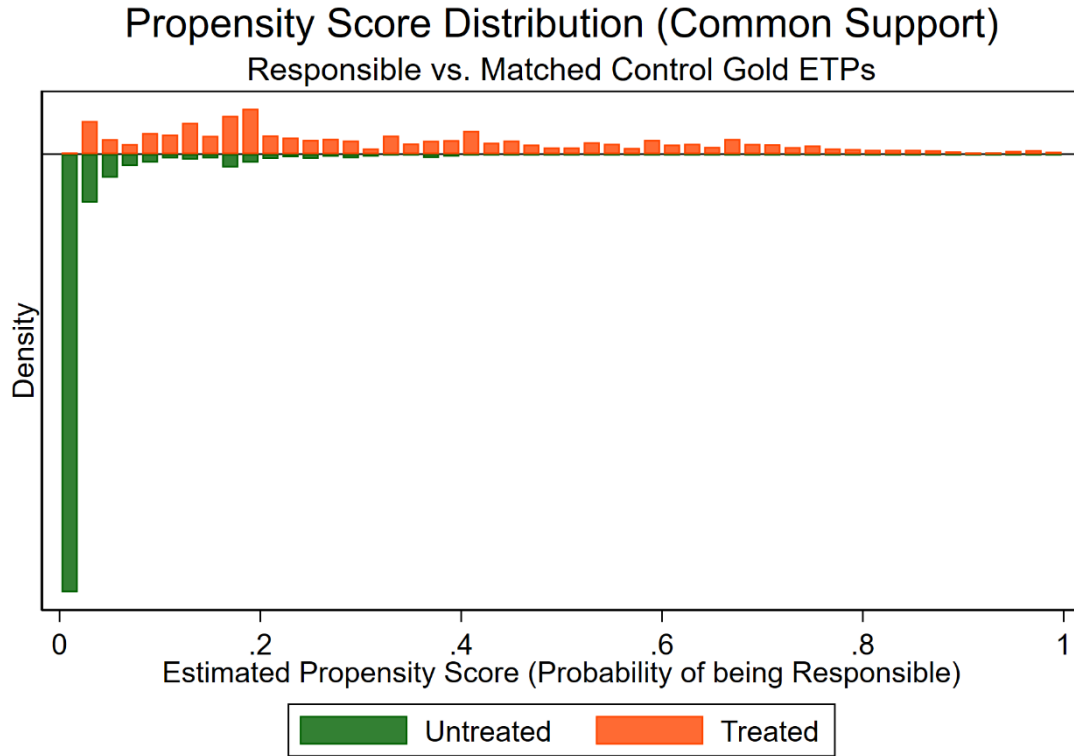
Upon estimating the propensity scores, we implement a k-Nearest Neighbor matching algorithm ($k=2$) with replacement. To ensure the quality of the matches and prevent bad matches between dissimilar funds, we impose a caliper of 0.05 standard deviations of the propensity score. This restrictive threshold ensures that each Responsible Gold ETP is compared against conventional counterparts with nearly identical structural profiles. By restricting our analysis to the region of common support, we effectively isolate the Responsible mandate effect from other confounding fund-level characteristics.

5.2 Matching Quality and Balance

To evaluate the reliability of the PSM procedure, we perform a series of diagnostic tests to ensure that the “Responsible” and “Matched Control” groups are comparable.

First, we examine the common support condition. Figure 4 presents the density distribution of the estimated propensity scores for both groups. The substantial overlap between the treated (Responsible) and untreated (Conventional) distributions indicates that the common support requirement is satisfied. While a significant portion of conventional funds with low propensity scores was correctly excluded by the algorithm, the remaining matched sample provides a robust basis for comparison across the entire range of probabilities.

Figure 4.



Second, we assess the balancing property of the covariates. A successful matching procedure should eliminate any statistically significant differences between the two groups' structural characteristics. Table 6 displays the means and differences for the matched sample.

Table 6. Matched sample

Variable	(1) Conventional (Control)	(2) Responsible (Treated)	(3) Difference	(4) % Bias Reduction
Log Market Value	3.977	3.597	-0.380	89.2%
Log Asset Age	5.631	6.154	0.523	58.2%
Total Expense Ratio	0.218	0.245	0.027	20.3%
Observations	2,738	2,687	5,425	

Note: This table reports covariate balance for the full matched sample obtained using Propensity Score Matching (1:2 nearest-neighbor matching with replacement; caliper 0.05). The larger number of Responsible Gold observations relative to Table 1 (2,687 vs. 1,340) arises because the propensity score is estimated using only the structural fund characteristics included in the matching model (log market value, log asset age, TER, and region), rather than the broader set of variables required for the baseline regressions reported in Table 1, Panel C. The larger control sample reflects the use of the maximum available historical information for each conventional fund (extending back to 2004 for the earliest LBMA-compliant vehicles) to maximize the pool of potential structural matches. The regression reported in Table 7, Column (2), further restricts the matched sample to the 2020-2025 estimation period used throughout the paper, reducing the effective sample to 1,173 observations.

As shown in Table 6, the matching procedure successfully minimizes the structural differences between groups. While the large sample size leads to statistically

significant t -tests, the economic magnitude of these differences is substantially reduced. Specifically, the matching achieves a standardized bias reduction of 89.2% for Log Market Value and 58.2% for Log Asset Age, and 20.3% for the Total Expense Ratio. Following Rosenbaum and Rubin (1985), this significant reduction in standardized bias confirms that the Responsible ETPs are compared against “statistical twins” from the conventional universe, effectively mitigating selection bias.

5.3 Robustness Results

The final stage of our robustness analysis involves re-estimating our baseline model using the matched sample derived from the PSM procedure. This approach aims to provide a more refined identification of the responsible mandate effect by comparing funds with nearly identical structural characteristics. Table 7 presents the results, comparing the full sample estimates (Column 1) with the matched sample (Column 2). To ensure the highest quality of the matching process by maximizing the pool of potential twins, the propensity score is estimated using the full historical data of the ETPs (from 2004 to 2025).

Table 7. Resilience during Volatility Shocks: PSM Robustness

	(1) Full Sample	(2) Matched Sample
Equity Volatility (VIX)	-0.032*** (0.006)	0.060 (0.090)
RESPONSIBLE $\times$ Equity Volatility (VIX)	0.008 (0.038)	-0.083 (0.098)
Bond Volatility (MOVE)	-0.041*** (0.007)	-0.128*** (0.029)
RESPONSIBLE $\times$ Bond Volatility (MOVE)	0.058*** (0.012)	0.150*** (0.032)
Lagged Spread	0.007 (0.018)	0.022* (0.011)
Log Market Value	0.093*** (0.018)	0.205 (0.142)
Log Trading Volume	-0.046*** (0.007)	-0.022 (0.013)
Observations	25928	1173
Adj. R-squared	0.050	0.022

Note: This table reports the re-estimation of the safe-haven baseline model (Table 3, Model 2) using a matched sample derived from a Propensity Score Matching (PSM) procedure. The dependent variable is the daily log-return of the ETP. Column (1) presents the full 2020-2025 estimation sample, restricted to the same period as Table 3, and is included here to confirm that the two specifications are directly comparable; accordingly, Column (1) exactly replicates the coefficients, standard errors, and observation count reported in Table 3, Model 2. Column (2) shows the results for the matched sample described in Table 6 (1-to-2 PSM with replacement), further restricted to the 2020-2025 estimation window (the period in which the Responsible Gold category exists and can be meaningfully compared against the LBMA-compliant benchmark, and which defines the estimation sample used throughout the paper) and to observations with non-missing values in all regression covariates. This reduces the sample from the 5,425 fund-day observations reported in Table 6 to 1,173. Standard errors are clustered at the fund level and reported in parentheses. LBMA-compliant is the baseline category. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

The results in Table 7 support our second hypothesis (H2). In the matched sample, we observe a clear decoupling between gold ETP segments during market distress. Specifically, the interaction coefficient for Responsible $\times$ Bond Volatility (MOVE) is positive and highly significant (0.150, $p < 0.01$) in the matched sample, showing a substantial increase in magnitude compared to the full sample (0.058). This indicates that, when comparing funds of similar size and age, Responsible Gold ETPs exhibit a significantly stronger "Safe Haven" premium than their conventional counterparts during fixed-income market stress.

Regarding equity market distress, the interaction term for Responsible $\times$ VIX (-0.083) remains statistically insignificant in the matched sample, consistent with the full sample results. These results suggest that standard bullion vehicles are more sensitive to liquidity shocks or capital reallocation during systemic distress, while the Responsible Gold segment shows a superior structural resilience.

The fact that the MOVE interaction remains robust and increases in magnitude in Column 2, despite the smaller matched sample size (1,173 observations), suggests that standard bullion vehicles are vulnerable to trading slowdowns or capital outflows during systemic shocks once fund size and age biases are controlled for. In contrast, the Responsible Gold segment exhibits properties of a stabilizing buffer, suggesting that the explicit sustainability mandate mitigates the downward pressure seen in the LBMA benchmark during periods of severe market distress.

5.4 Alternative Liquidity Proxy: Price Impact Analysis

To ensure that our results are not sensitive to the specific liquidity metric employed, we replicate the primary analysis using the Amihud (2002) illiquidity ratio. While the bid-ask spread primarily captures transaction costs, the Amihud ratio assesses market depth by measuring price impact per unit of volume. This distinction is crucial to confirm that the resilience of Responsible Gold reflects genuine market depth rather than just spread mechanics.

Table 8 presents the results of this multi-dimensional robustness check. To address potential concerns regarding multicollinearity or mechanical correlations between the Amihud ratio and its components (returns and volume), we calculated the Variance Inflation Factors (VIF). The mean VIF of 1.18, with the lagged Amihud ratio showing a VIF of 1.04, suggests that the model is econometrically sound and the safe-haven coefficients remain unbiased. Overall, the estimates provide a nuanced validation of our findings. In the first column (H1), the negative coefficient of lagged trading volume on the Amihud ratio (-9.232, $p < 0.10$) validates that higher trading activity enhances market depth. A key insight from Column 4 (H3) is that, unlike the bid-ask spread results, the Responsible Gold dummy shows no significant price-impact penalty. This suggests that the "Price of Ethics" acts primarily as a transaction-cost friction rather than a structural deficiency in market depth.

The most prominent results appear in the analysis of the safe-haven mechanism (H2). Looking at Columns 2 and 3, Responsible Gold exhibits a distinct behavior

during periods of market stress. Under elevated bond market volatility (MOVE shocks), these assets not only achieve significant positive returns (0.063, $p < 0.01$) but simultaneously experience a notable surge in trading volume (0.11, $p < 0.05$). This dual response is consistent with a potential capital reallocation process toward ethical mandates during periods of financial uncertainty, contrasting with the general market sensitivity where MOVE shocks typically reduce returns (-0.042, $p < 0.01$).

Finally, the results for the maturation stage (H3) confirm the persistence of illiquidity, captured by the lagged Amihud ratio (0.181 in col 1; 0.225 in col 4), as a highly significant factor across all segments. Collectively, this robustness test demonstrates that while investors may face higher entry costs, the market depth and safe-haven properties of Responsible Gold are structurally robust phenomena that hold across different measures of market friction.

Table 8. Robustness Check: Multi-dimensional Analysis (Amihud Proxy)

Dependent	H1: Attention Amihud Ratio	H2: Returns Daily Returns	H2: Volume Log Volume	H3: Maturation Amihud Ratio
Log Trading Volume lagged	-9.232* (4.868)			-9.209 (7.305)
Log Trading Volume		-0.047*** (0.007)		
Lagged Amihud (t-1)	0.181*** (0.043)	0.000 (0.000)		0.225*** (0.039)
Amihud Ratio			-0.000** (0.000)	
Responsible Gold (Dummy)	-34.947 (114.971)	0.000 (.)	-1.748* (0.943)	3.521 (115.395)
Attention: Gold	1.740 (6.496)			2.877 (6.187)
Responsible x Attention Gold	-87.893 (54.720)			-56.705 (43.966)
Attention: ESG	-11.498 (17.583)			-10.237 (15.892)
Responsible x Attention ESG	347.507* (176.519)			194.874 (153.869)
Log Market Value	2.231 (5.052)	0.099*** (0.024)	0.685*** (0.131)	0.452 (5.573)
Log Asset Age	-13.359 (10.845)		0.085 (0.257)	-17.164 (17.519)
Total Expense Ratio (TER)	-86.342 (64.945)		-0.511 (2.813)	-71.753 (86.226)
Financial Stress (VIX)	-0.703	-0.034***	0.008**	-1.190

	(0.951)	(0.005)	(0.004)	(1.926)
Bond Volatility (MOVE)	-0.321	-0.042***	0.012**	0.292
	(1.640)	(0.007)	(0.005)	(1.594)
Geopolitical Threat	-2.315	-0.014***	0.010*	-3.766
	(1.574)	(0.004)	(0.005)	(3.955)
Geopolitical Acts	-0.821	0.031***	-0.005	-0.111
	(2.468)	(0.004)	(0.005)	(2.729)
Responsible x VIX Shock		-0.014	-0.006	
		(0.026)	(0.015)	
Responsible x MOVE Shock		0.063***	0.110**	
		(0.009)	(0.042)	
Responsible x GPR Threat		-0.004	-0.010	
		(0.046)	(0.041)	
Responsible x GPR Acts		-0.025	-0.011	
		(0.030)	(0.032)	
Daily Return (%)			-0.033***	
			(0.009)	
Region FE	Yes		Yes	Yes
Month FE	Yes			Yes
Asset FE		Yes		
Year-Month FE		Yes	Yes	
Adj. R-squared	0.166	0.053	0.831	0.170
Observations	24764	24764	26117	13084

Note: Standard errors clustered at the asset level are reported in parentheses. LBMA-compliant is the baseline category. This table presents a multi-dimensional robustness analysis using the Amihud (2002) illiquidity ratio as the primary proxy for price impact in Columns (1) and (4). The Amihud ratio is calculated as the daily absolute return divided by the dollar trading volume (multiplied by 10^9 for scale); lower values indicate greater market depth and higher liquidity. Column (2) employs Daily Logarithmic Returns (%) to test the price dimension of the safe-haven hypothesis (H2), while Column (3) uses Log Trading Volume to capture trading activity associated with the hypothesized capital reallocation. In Column (2), Asset Fixed Effects absorb all time-invariant asset characteristics. Consequently, the main effect of Responsible Gold is perfectly collinear with the Asset Fixed Effects and cannot be estimated separately. The reported value of 0.000 (.) therefore denotes an omitted coefficient rather than an estimated effect. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

5.5 Capital Flow Dynamics

To further assess whether the trading-volume response documented above is consistent with an underlying reallocation of capital, we examine capital flows. We proxy capital reallocation through trading volume rather than percentage-based fund flows. The standard net growth rate measure (Sirri & Tufano, 1998) is mechanically sensitive to fund size: as total net assets grow, identical absolute inflows produce smaller percentage flows. Given the significant scale differential between LBMA and Responsible segments (mean Market Value:\$8,477 vs. \$118 million), this denominator effect would confound any cross-sectional or temporal comparison of percentage flows. To verify that the volume-based evidence reflects

genuine capital dynamics, we compute dollar-denominated flows at a monthly frequency:

$$DollarFlow_{i,t} = MV_{i,t} - MV_{i,t-1}(1 + R_{i,t})$$

The results in Table 9 provide two critical insights. First, we document structural stability in capital attraction for both segments. As shown in Panels A and B, absolute inflows did not decline as the market matured; in fact, Responsible Gold ETPs maintained steady monthly inflows (+2.42 vs +3.12 USD mn, $p=0.822$), consistent with the view that any observed decline in percentage growth is largely a scaling artifact.

Second, the data reveals a divergence in how segments respond to fixed-income instability. Conventional LBMA ETPs (Panel A) exhibit insensitivity to bond market stress, with monthly flows remaining virtually unchanged. Conversely, the point estimates for Responsible ETPs show a substantial increase, with monthly inflows rising from 2.12 USD mn during calm periods to 6.28 USD mn during MOVE stress episodes ($p < 0.10$), consistent with an increase in capital reallocation during bond market distress. This divergent behavior is consistent with a possible flight-to-responsibility pattern, whereby Responsible Gold may attract relatively greater capital during periods of elevated bond-market stress. Taken together with the trading-volume evidence, these results provide suggestive support for the interpretation that periods of bond-market stress may be associated with capital reallocation toward Responsible Gold ETPs.

Table 9. Robustness Analysis: Monthly Dollar Flows
Panel A. LBMA Gold

	By Era			By MOVE		
	Inception	Mature	Diff.	Calm	Stress	Diff.
DollarFlow	32.082	88.407	-56.325	61.499	62.353	-.854
p-value (2-sided)			0.146			0.985
p-value (1-sided)			0.073			0.492

Panel B. Responsible Gold

	By Era			By MOVE		
	Inception	Mature	Diff.	Calm	Stress	Diff.
DollarFlow	2.416	3.118	-.7021	2.122	6.284	-4.162
p-value (2-sided)			0.822			0.131
p-value (1-sided)			0.411			0.065

Note: Dollar flows are reported in USD mn. Inception Era (2020–22), Maturation Era (2023–25). Stress periods are defined by the top quartile of the monthly standardized MOVE index. Both two-sided and one-sided p-values from the t-tests are reported.

5.6 Robustness to Individual Fund Exclusion

Given that the Responsible Gold ETP segment is currently composed of a relatively small number of specialized vehicles, a potential concern is whether our primary findings are driven by the idiosyncratic behavior of a single large or highly active fund. To address this, we conduct a leave-one-out sensitivity analysis, sequentially re-estimating our core models from Tables 3, 4, and 8 while excluding one

responsible ETP at a time. This procedure allows us to identify if any specific asset exerts a disproportionate influence on the aggregate coefficients.

Table 10 reports the results of this sensitivity test for the key interaction coefficients with bond market volatility (MOVE). The safe-haven return profile (Column 1, based on Table 3) is remarkably stable. The interaction coefficient for Responsible $\times$ MOVE remains positive and statistically significant at the 1% level across all five iterations. Even when the most influential funds are excluded, the responsible segment continues to exhibit a significant return buffer during bond market shocks. This suggests that price resilience is a structural characteristic of the entire responsible gold category rather than an anomaly driven by a specific vehicle.

Second, the capital reallocation volume (Columns 2 and 3, based on Tables 4 and 8) shows greater sensitivity to sample composition. While the volume surge remains significant or even strengthens when most funds are excluded (notably when excluding 2615KW, where the coefficient rises to 0.164), the exclusion of fund 2675P0 (Franklin Responsibly Sourced Gold ETF) leads to a loss of statistical significance in the volume interaction. This indicates that while the safe-haven return benefit is universal across the segment, the evidence of active capital reallocation is particularly sensitive to this specific vehicle and should therefore be interpreted with greater caution.

Overall, the leave-one-out analysis confirms the internal validity of our safe-haven results. While the magnitude of trading activity varies across individual ETPs, the core finding of a structural return resilience during systemic financial stress is robust to the exclusion of any single asset in the responsible gold universe.

Table 10. Leave-One-Out Sensitivity Analysis

	(1)	(2)	(3)
Excluded Fund	Returns (T3)	Volume (T4)	Amihud (T8)
Baseline (Full Sample)	0.058***	0.117**	0.110**
2675P0	0.043***	0.019	0.013
2729JD	0.064***	0.109**	0.094**
2810FQ	0.058***	0.117**	0.110**
96153Q	0.058***	0.117**	0.110**
2615KW	0.062***	0.164***	0.155***

Note: This table presents a "leave-one-out" sensitivity analysis to assess the robustness of the safe-haven effect across the responsible gold segment. Each row reports the estimated interaction coefficient between the Responsible Gold dummy and Bond Market Volatility (MOVE) after sequentially excluding one of the five responsible ETPs from the sample. The Baseline row provides the coefficients for the full sample for comparison. Column (1) replicates the safe-haven return model from Table 3 (col. 2); Column (2) replicates the trading volume model from Table 4 (col. 3); and Column (3) replicates the volume model from Table 8 (Col. 3), which incorporates the Amihud price-impact ratio as an additional control. The results confirm that the safe-haven return profile is a structural feature of the entire category, while the magnitude of capital reallocation is more sensitive to individual fund composition. Standard errors clustered at the asset level.

*** p<0.01, ** p<0.05, * p<0.10.

5.7 Randomization Inference

A potential limitation of our empirical design is the small number of Responsible Gold ETPs (N = 5) relative to the broader sample. To ensure that our findings are not

primarily driven by sample composition or a few influential treated units, we implement Randomization Inference (RI) following the estimation procedure developed by Heß (2017). RI is particularly well suited to this setting because it relies on the assignment mechanism rather than large-sample asymptotic approximations, making it a robust alternative in contexts with a limited number of treated units (MacKinnon and Webb, 2020).

We perform 1,000 random permutations in which the “Responsible” label is reassigned across the universe of 34 gold funds while preserving the number of treated units. For each permutation, we re-estimate the specifications reported in Tables 2-5 to construct the empirical distribution of placebo coefficients. Table 11 summarizes the resulting randomization p-values alongside the corresponding cluster-robust p-values reported in Tables 2-5.

Table 11. Randomization Inference

Coefficient	Estimate	Cluster-robust p-value	Randomization p-value
Table 2 - Responsible Gold (Baseline Model)	0.131	0.001	0.027
Table 2 - Responsible x Attention Gold	0.073	< 0.01	0.005
Table 2 - Responsible x Attention ESG	-0.110	0.114	0.124
Table 3 - Responsible x MOVE	0.058	< 0.01	0.011
Table 4 - Responsible x MOVE	0.117	0.011	0.002
Table 5 - Responsible x Attention Gold (Inception)	-0.082	0.01	0.089
Table 5 - Responsible x Attention Gold (Maturation)	0.074	< 0.01	0.002
Table 5 - Responsible x Attention ESG (Inception)	0.150	< 0.01	0.473
Table 5 - Responsible x Attention ESG (Maturation)	-0.155	< 0.01	0.003

Note: Randomization p-values are based on 1,000 permutations of the Responsible label, clustered at the fund level, following the same procedure applied to Tables 2-5.

The randomization inference results broadly confirm the conclusions obtained from the conventional cluster-robust inference. The baseline Responsible Gold premium remains statistically significant (Table 2, Baseline Model; randomization $p = 0.027$), while the safe-haven results are likewise confirmed, with both the return interaction (Table 3, randomization $p = 0.011$) and the trading-volume interaction (Table 4, randomization $p = 0.002$) remaining statistically significant. The latter result should nevertheless be interpreted with some caution, as the leave-one-out analysis reported in Section 5.6 shows that the trading-volume interaction is sensitive to the

exclusion of fund 2675P0. Thus, the randomization inference supports the full-sample association between Responsible status and trading activity during periods of elevated bond-market volatility, but does not eliminate the sensitivity of this result to individual fund composition.

For the attention models, the evidence is more nuanced. The Responsible $\times$ Attention Gold interaction remains statistically significant both in the pooled specification (Table 2, randomization $p = 0.005$) and in the Maturation subsample (Table 5, randomization $p = 0.002$). Likewise, the Responsible $\times$ Attention ESG interaction in the Maturation subsample remains statistically significant (randomization $p = 0.003$). These results support our conclusion that thematic and gold-specific attention are priced differently once the Responsible Gold segment reaches its mature phase. Importantly, these findings remain robust under this more conservative inference procedure.

By contrast, the Inception-period interactions receive more limited support under randomization inference. The Responsible $\times$ Attention Gold interaction is only marginally significant ($p = 0.089$), while the Responsible $\times$ Attention ESG interaction is no longer statistically significant ($p = 0.473$). Accordingly, the evidence for the Inception period is more tentative than that for the Maturation period, which remains strongly supported under randomization inference.

5.8 Limitations

While our empirical design incorporates several robust estimation techniques, we acknowledge certain limitations. First, the sample of Responsible Gold ETPs remains relatively small ($N = 5$), reflecting the nascent nature of this market segment. Although these ETPs contribute 1,340 fund-day observations to the baseline sample, compared with 31,029 observations from the 29 LBMA-compliant ETPs, the daily observations are concentrated in only five treated funds. Thus, the number of observations should not be interpreted as equivalent to a large number of independent treated units. The five Responsible ETPs also differ substantially in terms of fund size. As of December 31, 2025, their AUM ranges from USD 8 million to USD 1,573 million. This indicates considerable heterogeneity within the treatment group. The Responsible indicator therefore captures a common responsible-sourcing characteristic, but does not imply that the five treated ETPs are otherwise homogeneous. Differences in fund size and other fund-level characteristics may contribute to heterogeneity within the treated group. The Leave-One-Out analysis shows that the estimated safe-haven effect on returns remains significant when each Responsible ETP is excluded in turn, while the trading-volume response is more sensitive to the composition of the Responsible group. Randomization Inference provides additional evidence under an alternative inference procedure, although it also indicates more limited support for some interaction effects during the market's inception phase. These results suggest that the main findings are reasonably robust within the observed Responsible Gold segment, but should be interpreted with caution when considering their generalizability beyond the five ETPs in our sample.

Second, regarding potential survivorship bias, we have cross-referenced our sample with LSEG Eikon to ensure that all gold-backed ETPs active during the 2020-2025 period are included. Furthermore, our Propensity Score Matching (PSM) procedure utilizes a comprehensive database dating back to the inception of the gold ETP market in the early 2000s, incorporating funds that have since been delisted, liquidated, or merged. This ensures that our baseline comparisons are not biased toward surviving or high-performing funds.

Finally, while we focus on secondary market liquidity (bid-ask spreads and Amihud ratios), we note that primary market creation-redemption dynamics, driven by authorized participants, remain unobserved. Future research could explore these multi-layered liquidity structures as more granular data on institutional arbitrage activity becomes available.

6 Concluding Remarks and Policy Implications

Our analysis uncovers several key insights into the evolving liquidity-resilience dynamics of Responsible Gold ETPs and their distinction from conventional gold instruments. First, we document a consistent responsibility premium embedded in transaction-costs. Responsible Gold ETPs consistently exhibit wider bid-ask spreads than their standard counterparts, even after accounting for trading volume and size. This persistent liquidity penalty suggests that the responsible label is associated with additional frictions, consistent with the presence of a more segmented investor clientele as theorized in the sustainable finance literature, as well as the informational demands associated with traceability and sourcing transparency.

Second, we highlight the differentiated safe-haven behavior of responsible gold during episodes of market stress. While gold broadly hedges equity market turbulence, Responsible Gold ETPs display a comparatively stronger resilience to shocks originating in the bond market, particularly those captured by MOVE volatility shocks. This pattern remains robust when controlling for fund characteristics through Propensity Score Matching and when examining alternative liquidity measures such as the Amihud ratio. The coexistence of price resilience and heightened trading activity during bond market distress is consistent with a potential reallocation of capital toward responsible vehicles in the full sample. However, the leave-one-out analysis indicates that the volume response is sensitive to the composition of the Responsible segment, as the Responsible $\times$ MOVE coefficient becomes statistically insignificant when fund 2675P0 is excluded. Accordingly, the evidence for capital reallocation should be interpreted as suggestive rather than as a general characteristic of the Responsible Gold segment. The supporting evidence from price impact metrics is consistent with the view that this resilience is supported by adequate market depth, providing indirect evidence of a more stable investor base consistent with the behavior documented for ESG-mandated assets capable of absorbing fixed-income volatility.

Third, our findings reveal that market maturation is associated with a significant evolution in the liquidity structure of responsible gold. In the early stages of the market, increased ESG attention appears to have widened spreads relative to the benchmark, reflecting novelty effects and informational uncertainty. As the market matures, however, thematic attention is associated with spread compression, even as the responsibility premium becomes more firmly priced as a structural cost. At the same time, market depth exhibits signs of convergence across responsible and conventional segments, indicating that while transaction-cost differentials persist, they are partially offset by the asset's relative resilience during bond market stress and its comparable price impact profile in mature trading environments.

These results carry important implications for investors and regulators. For portfolio managers, Responsible Gold ETPs exhibit characteristics that distinguish them from purely liquidity driven instruments, exhibiting characteristics of specialized defensive assets that trade off higher transaction-costs for enhanced resilience to fixed-income volatility. Regulators, meanwhile, may find value in monitoring how responsible designations interact with market structure as these products continue to evolve. Future research could explore whether similar maturation patterns emerge in other responsible commodity markets or whether the liquidity resilience trade-off documented here is unique to the institutional and historical context of gold.

References

- Amihud, Y. (2002). Illiquidity and stock returns: cross-section and time-series effects. *Journal of Financial Markets*, 5(1), 31–56. [https://doi.org/10.1016/S1386-4181\(01\)00024-6](https://doi.org/10.1016/S1386-4181(01)00024-6)
- Angelini, E., Foglia, M., Ortolano, A., & Leone, M. (2018). The “Donald” and the market: Is there a cointegration? *Research in International Business and Finance*, 45, 30–37. <https://doi.org/10.1016/j.ribaf.2017.07.129>
- Awartani, B., Hussain, S. M., & Virk, N. (2024). How do the gold intra-day returns and volatility react to monetary policy shocks? *International Review of Financial Analysis*, 95. <https://doi.org/10.1016/j.irfa.2024.103486>
- Bae, K., & Kim, D. (2020). Liquidity risk and exchange-traded fund returns, variances, and tracking errors. *Journal of Financial Economics*, 138(1), 222–253. <https://doi.org/10.1016/J.JFINECO.2019.02.012>
- Barber, B. M., & Odean, T. (2008). All That Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors. *The Review of Financial Studies*, 21(2), 785–818. <https://doi.org/10.1093/rfs/hhm079>
- Barro, R. J., & Misra, S. (2016). Gold Returns. *The Economic Journal*, 126(594), 1293–1317. <https://doi.org/10.1111/ecoj.12274>

- Basher, S. A., & Sadorsky, P. (2016). Hedging emerging market stock prices with oil, gold, VIX, and bonds: A comparison between DCC, ADCC and GO-GARCH. *Energy Economics*, 54, 235–247. <https://doi.org/10.1016/j.eneco.2015.11.022>
- Baur, D. G., & Lucey, B. M. (2010). Is gold a hedge or a safe haven? An analysis of stocks, bonds and gold. *Financial Review*, 45(2), 217–229. <https://doi.org/10.1111/j.1540-6288.2010.00244.x>
- Baur, D. G., & McDermott, T. K. (2010). Is gold a safe haven? International evidence. *Journal of Banking and Finance*, 34(8), 1886–1898. <https://doi.org/10.1016/j.jbankfin.2009.12.008>
- Ben-David, I., Franzoni, F., & Moussawi, R. (2017). Exchange-Traded Funds. *Annual Review of Financial Economics*, 9(1), <https://doi.org/10.1146/annurev-financial-110716-032538>
- Ben-David, I., Franzoni, F., & Moussawi, R. (2018). Do ETFs Increase Volatility? *The Journal of Finance*, 73(6), 2471–2535. <https://doi.org/10.1111/jofi.12727>
- Bilgin, M. H., Gozgor, G., Lau, C. K. M., & Sheng, X. (2018). The effects of uncertainty measures on the price of gold. *International Review of Financial Analysis*, 58, 1–7. <https://doi.org/10.1016/j.irfa.2018.03.009>
- Bollen, N. P. B. (2007). Mutual Fund Attributes and Investor Behavior. *Journal of Financial and Quantitative Analysis*, 42(3), 683–708. <https://doi.org/DOI:10.1017/S0022109000004142>
- Caldara, D., & Iacoviello, M. (2022). Measuring Geopolitical Risk. *American Economic Review*, 112(4), 1194–1225. <https://doi.org/10.1257/aer.20191823>
- Clifford, C. P., Fulkerson, J. A., & Jordan, B. D. (2014). What drives ETF flows? *Financial Review*, 49(3), 619–642. <https://doi.org/10.1111/FIRE.12049>
- Correia, Sergio. 2017. “Linear Models with High-Dimensional Fixed Effects: An Efficient and Feasible Estimator” Working Paper. <http://scorreia.com/research/hdfe.pdf>
- Da, Z., Engelberg, J., & Gao, P. (2011). In Search of Attention. *Journal of Finance*, 66(5), 1461–1499. <https://doi.org/10.1111/j.1540-6261.2011.01679.x>
- de Pessemier, J., Palmberg, J., Chacko, K., Gopaul, K., Street, L., Salim, M., Jia, R., Burnette, T., Artigas, J. C., Reade, J., & Cavatoni, J. (2026). *Gold as a strategic asset (2026 edition)*. World Gold Council. <https://www.gold.org/goldhub/research/relevance-of-gold-as-a-strategic-asset>
- Ding, Q., Huang, J., Gao, W., & Zhang, H. (2022). Does political risk matter for gold market fluctuations? A structural VAR analysis. *Research in International Business and Finance*, 60. <https://doi.org/10.1016/j.ribaf.2022.101618>

Echaust, K., & Just, M. (2022). Is gold still a safe haven for stock markets? New insights through the tail thickness of portfolio return distributions. *Research in International Business and Finance*, 63.

<https://doi.org/10.1016/j.ribaf.2022.101788>

Erb, Claude B. and Harvey, Campbell R. (2025), *Understanding Gold*. Available at <http://dx.doi.org/10.2139/ssrn.5525138>

Friede, G., Busch, T., & Bassen, A. (2015). ESG and financial performance: aggregated evidence from more than 2000 empirical studies. *Journal of Sustainable Finance and Investment*, 5(4), 210–233.

<https://doi.org/10.1080/20430795.2015.1118917>

Glosten, L. R., & Milgrom, P. R. (1985). Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics*, 14(1), 71–100.

Harvey, C. R. (2025), Gold and Bitcoin. Available at SSRN:

<https://ssrn.com/abstract=5530719>

Harvey, C. R., Hoyle, E., Rattray, S., Sargaison, M., Taylor, D., & Hemert, O. van. (2019). The Best of Strategies for the Worst of Times: Can Portfolios Be Crisis Proofed? *The Journal of Portfolio Management*, 45(5), 7–28.

<https://doi.org/10.3905/jpm.2019.45.5.007>

Heß, S. (2017). Randomization Inference with Stata: A Guide and Software. *The Stata Journal*, 17(3), 630–651.

He, Z., O'Connor, F., & Thijssen, J. (2018). Is gold a Sometime Safe Haven or an Always Hedge for equity investors? A Markov-Switching CAPM approach for US and UK stock indices. *International Review of Financial Analysis*, 60, 30–37.

<https://doi.org/10.1016/j.irfa.2018.08.010>

Hood, M., & Malik, F. (2013). Is gold the best hedge and a safe haven under changing stock market volatility? *Review of Financial Economics*, 22(2), 47–52.

<https://doi.org/10.1016/j.rfe.2013.03.001>

Kim, J., & Yang, W. (2025). Environmental, Social, and Governance (ESG) Research: A Systematic Review of Recent Trends (2020–2024). *Sustainable Development*, 1–29. <https://doi.org/10.1002/sd.70370>

Lehnert, T. (2022). Flight-to-safety and retail investor behavior. *International Review of Financial Analysis*, 81. <https://doi.org/10.1016/j.irfa.2022.102142>

Li, X., Guo, Q., Liang, C., & Umar, M. (2023). Forecasting gold volatility with geopolitical risk indices. *Research in International Business and Finance*, 64.

<https://doi.org/10.1016/j.ribaf.2022.101857>

- Liebi, L. J. (2020). The effect of ETFs on financial markets: a literature review. *Financial Markets and Portfolio Management*, 34(2), 165–178. <https://doi.org/10.1007/S11408-020-00349-1>
- Lo, A. W. (2004). The Adaptive Markets Hypothesis. *The Journal of Portfolio Management*, 30(5), 15–29. <https://doi.org/10.3905/jpm.2004.442611>
- MacKinnon, J. G., & Webb, M. D. (2020). Randomization inference for difference-in-differences with few treated clusters. *Journal of Econometrics*, 218(2), 435–450. <https://doi.org/10.1016/j.jeconom.2020.04.024>
- Merton, R. C. (1987). A Simple Model of Capital Market Equilibrium with Incomplete Information. *The Journal of Finance*, 42(3), 483–510.
- Michaud, Richard O. and Michaud, Robert and Pulvermacher, Katharine, Gold as a Strategic Asset (December 1, 2011). World Gold Council, 10 Old Bailey, London EC4M 7NG, UK, 2006, Available at SSRN: <https://ssrn.com/abstract=2402862>
- Oakley, P. (2022). Searching for pure gold: The impact of ethical gold sourcing certification programmes in the UK and Switzerland. *Environmental Science and Policy*, 132, 101–108. <https://doi.org/10.1016/j.envsci.2022.02.016>
- Parnes, D., & Parnes, S. S. (2025). Hedging geopolitical risks with diverse commodities. *International Review of Financial Analysis*, 102. <https://doi.org/10.1016/j.irfa.2025.104129>
- Pattnaik, D., Hassan, M. K., DSouza, A., & Ashraf, A. (2023). Investment in gold: A bibliometric review and agenda for future research. In *Research in International Business and Finance* (Vol. 64). Elsevier Ltd. <https://doi.org/10.1016/j.ribaf.2022.101854>
- Pástor, L., Stambaugh, R. F., & Taylor, L. A. (2021). Sustainable investing in equilibrium. *Journal of Financial Economics*, 142(2), 550–571. <https://doi.org/10.1016/j.jfineco.2020.12.011>
- Piñeiro-Chousa, J., López-Cabarcos, M. Á., Pérez-Pico, A. M., & Ribeiro-Navarrete, B. (2018). Does social network sentiment influence the relationship between the S&P 500 and gold returns? *International Review of Financial Analysis*, 57, 57–64. <https://doi.org/10.1016/j.irfa.2018.02.005>
- Renneboog, L., Ter Horst, J., & Zhang, C. (2008). Socially responsible investments: Institutional aspects, performance, and investor behavior. *Journal of Banking and Finance*, 32(9), 1723–1742. <https://doi.org/10.1016/j.jbankfin.2007.12.039>
- Rosenbaum, P. R., & Rubin, D. B. (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika*, 70(1), 41–55. <https://doi.org/10.2307/2335942>

- Sirri, E. R., & Tufano, P. (1998). Costly Search and Mutual Fund Flows. *The Journal of Finance*, 53(5), 1589–1622. <https://doi.org/10.1111/0022-1082.00066>
- Stoll, H. R. (1989). Inferring the Components of the Bid-Ask Spread: Theory and Empirical Tests. *The Journal of Finance*, 44(1), 115–134.
- 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 in International Business and Finance*, 67. <https://doi.org/10.1016/j.ribaf.2023.102125>
- Yang, N., Fernandez-Perez, A., & Indriawan, I. (2024). Spillover between investor sentiment and volatility: The role of social media. *International Review of Financial Analysis*, 96. <https://doi.org/10.1016/j.irfa.2024.103643>
- Younis, I., Naeem, M. A., Shah, W. U., & Tang, X. (2025). Inter- and intra-connectedness between energy, gold, Bitcoin, and Gulf cooperation council stock markets: New evidence from various financial crises. *Research in International Business and Finance*, 73. <https://doi.org/10.1016/j.ribaf.2024.102548>
- Yunus, N. (2020). Time-varying linkages among gold, stocks, bonds and real estate. *Quarterly Review of Economics and Finance*, 77, 165–185. <https://doi.org/10.1016/j.qref.2020.01.015>
- Zhu, Y., Fan, J., & Tucker, J. (2018). The impact of monetary policy on gold price dynamics. *Research in International Business and Finance*, 44, 319–331. <https://doi.org/10.1016/j.ribaf.2017.07.100>

Appendix A

TABLE A.1 Variable Definitions

Variable	Type	Description / Construction	Source
Spread	Dependent /Control	Relative Bid-Ask Spread: $(Ask_{it} - Bid_{it})/Price_{it} \times 100$.	LSEG Datastream
Daily Return	Dependent	Logarithmic daily return: $\ln(Price_t/Price_{t-1}) \times 100$.	LSEG Datastream
Responsible	Indicator	Dummy = 1 if the ETP is explicitly marketed as “Responsible”	LSEG Datastream
LBMA Gold	Indicator	Dummy = 1 if the ETP is LBMA-compliant but not explicitly marketed as “Responsible”.	LSEG Datastream
MOVE Shock	Shock	Daily change of ICE BofA MOVE Index (Bond Volatility). Standardized (Z-score)	LSEG Datastream
VIX Shock	Shock	Daily change of the CBOE Volatility Index (Equity Volatility). Standardized (Z-score)	FRED
GPR Threats/Acts	Shock	Daily change indices of geopolitical risk (Threats and Acts) as per Caldara & Iacoviello (2022). Standardized (Z-score)	authors' website
Attention: ESG/GOLD	Attention	Natural logarithm of one plus monthly Google Search Volume (GSV) for terms “Sustainable Gold” and “ESG Investing”. Standardized (Z-score)	Google Trends
Log MV	Control	Natural logarithm of the Fund's Total Market Value in USD.	LSEG Datastream
Log Age	Control	Natural logarithm of days since the ETP's inception (Seasoning).	LSEG Datastream
TER	Control	Total Expense Ratio (annualized %). Missing values supplemented via hand-collection.	LSEG / Hand-collected

Log Volume	Dep./Ctrl	Natural logarithm of the daily dollar trading volume. (Note: Excluded as control in Eq. 3).	LSEG Datastream
Amihud	Robustnetss	Price Impact ratio: $Amihud_{i,t} = \left(\frac{ R_{i,t} }{Volume_{i,t}} \right) \times 10^9$ It measures the absolute daily return per dollar of trading volume, capturing market depth as per Amihud (2002).	LSEG Datastream

Note: Given the mixed-frequency nature of our dataset, monthly indicators (Attention: ESG/GOLD) are forward-filled to a daily frequency to ensure temporal alignment with ETP-level liquidity and return series. This approach maintains the monthly value constant across all trading days within the corresponding month. Similarly, the Total Expense Ratio (TER) is treated as a time-invariant characteristic using the most recent available data for each ETP. This approach ensures temporal alignment with the daily liquidity and return series while accounting for the structural cost differences between funds.

TABLE A.2 ETPs List

Datastream Code	ISIN	ETP Name	Market	Original Currency	Country	AUM (USD\$mn)	Gold type
9449PC	FR0013416716	Amundi Physical Gold ETC	Euronext Paris	EUR	France	11,778	LBMA
2802G5	CA05611P1045	BMO Gold Bullion Series Units ETF USD	Toronto SE	USD	Canada	25	LBMA
87677Q	DE000EWGOLD1	Boerse Stuttgart Commodities ETC	Boerse Stuttgart	EUR	Germany	971	LBMA
2886FR	DE000EWG4TR6	Boerse Stuttgart Commodities Gold IHS 2025 ETC	XETRA	EUR	Germany	4	LBMA
2886FQ	DE000EWG4CR2	Boerse Stuttgart Commodities Gold IHS 2025 ETC	XETRA	EUR	Germany	13	LBMA
9170J2	DE000EWG2LD7	Boerse Stuttgart Securities Gold IHS 2017(17/Und) ETC	Boerse Stuttgart	EUR	Germany	4,202	LBMA
2577VY	DE000TMGOLD6	Boerse Stuttgart Securities Gold IHS 2021(22/Und) ETC	Boerse Stuttgart	EUR	Germany	159	LBMA
9926N1	CA12562B2093	CI Gold Bullion Fund ETF Series Units	Toronto SE	USD	Canada	422	LBMA
68831J	DE000A0S9GB0	Deutsche Boerse Commodities Xetra-Gold ETC	Boerse Frankfurt	EUR	Germany	23,944	LBMA
86785H	GB00B00FHZ82	Gold Bullion Securities ETC	London SE	GBP	UK	4,389	LBMA
9325FA	US38150K1034	Goldman Sachs Physical Gold ETF	CBOE	USD	US	2,543	LBMA
91406W	US38748G1013	GraniteShares Gold Trust	NYSE ARCA	USD	US	1,492	LBMA
69930J	IE00B579F325	Invesco Physical Gold ETC	London SE	USD	UK	28,448	LBMA
75736W	IE00B4ND3602	Ishares Physical Gold ETC	London SE	USD	UK	33,680	LBMA
9372M7	CA74642E3068	Purpose Gold Bullion Fund ETF	Toronto SE	CAD	Canada	574	LBMA
86548D	CH0134034930	Raiffeisen ETF - Solid Gold A USD	Six Swiss Exchange	USD	Switzerland	816	LBMA

779809	CA7799211056	Royal Canadian Mint Canadian Gold Reserves	Toronto SE	CAD	Canada	792	LBMA
9313LL	US98149E3036	SPDR Gold MiniShares Trust	NYSE ARCA	USD	US	25,288	LBMA
29793W	US78463V1070	SPDR Gold Shares	NYSE ARCA	USD	US	148,228	LBMA
32961L	CH0139101593	Swisscanto (CH) Gold ETF EA CHF ETF	Six Swiss Exchange	CHF	Switzerland	21,169	LBMA
68706X	CH0106027144	UBS Gold hEUR ETF EUR dis	Six Swiss Exchange	EUR	Switzerland	1,021	LBMA
8897P6	US9210781012	VanEck Merk Gold ETF	NYSE ARCA	USD	US	2,459	LBMA
99130Z	JE00BN2CJ301	WisdomTree Core Physical Gold ETC	London SE	USD	UK	1,961	LBMA
68542V	JE00B588CD74	WisdomTree Physical Swiss Gold ETC	London SE	USD	UK	5,235	LBMA
69679K	GB00B5840F36	Xtrackers Physical Gold Etc	London SE	USD	Germany	553	LBMA
67967M	US00326A1043	abrdn Physical Gold Shares ETF	NYSE ARCA	USD	US	7,333	LBMA
68128Q	CH0104136236	iShares Gold ETF (CH)	Six Swiss Exchange	USD	Switzerland	1,472	LBMA
30218C	US4642852044	iShares Gold Trust	NYSE ARCA	USD	US	68,377	LBMA
2584Q0	US46436F1030	iShares Gold Trust Micro	NYSE ARCA	USD	US	6,029	LBMA
2675P0	US35473M1053	Franklin Responsibly Sourced Gold ETF	NYSE ARCA	USD	US	465	RESPONSIBLE
2615KW	CH1122756724	Raiffeisen ETF - Solid Gold Resp Source&Trcb A CHF	Six Swiss Exchange	CHF	Switzerland	88	RESPONSIBLE
2810FQ	XS2792094604	SMO Physical Gold ETC	London SE	USD	UK	8	RESPONSIBLE
96153Q	XS2115336336	TRM Phys Gold ETC ETP	London SE	USD	Ireland	1,573	RESPONSIBLE
2729JD	CH1233056329	UBS Carbon Compensated Gold ETF USD acc	Six Swiss Exchange	USD	Switzerland	48	RESPONSIBLE