# Does divestment move risk? A null result from the world's largest sovereign wealth fund

Simen Guttormsen[a,*]

[a] School of Economics and Business, Norwegian University of Life Sciences, Ås, Norway

ARTICLE INFO

*Keywords*:
Divestment
ESG
Sovereign wealth fund
Volatility
Event study
GARCH
NBIM

ABSTRACT

Theory predicts that institutional divestment raises firm-level volatility. This paper tests the prediction using 181 exclusions by Norway's Government Pension Fund Global, the world's largest sovereign wealth fund. GARCH(1,1) variance ratio tests show no effect at announcement or during the preceding investigation. *Product-based* exclusions (tobacco, weapons, coal) carry no underlying scandal, isolating the divestment channel; *conduct-based* exclusions bundle divestment with an ESG controversy, confounding the two effects. Panel difference-in-differences estimates are small and insignificant, and equivalence tests tightly bound the true effect around zero. Combined with the cost-of-capital null of Berk and Van Binsbergen (2025), these findings indicate that neither theoretical channel—cost of capital or volatility—is empirically active.

## 1. Introduction

Does institutional divestment increase the volatility of excluded firms? The theory suggests that it does. Merton et al. (1987) argued that when the investor base for a security narrows, each remaining investor holds a larger share, carrying more idiosyncratic risk, requiring a higher rate of return. Heinkel et al. (2001) extended Merton's reasoning to include an ethical boycott where ethical investors refuse to purchase shares in polluting firms. This makes the remaining shareholders responsible for absorbing the risk. As a result, the cost of capital increases. Hong and Kacperczyk (2009) empirically demonstrated that norm-constrained institutions systematically exclude sin stocks, despite the stocks producing approximately 250 bps per year of excess returns. The intuition is straightforward. If a large institutional investor reduces its holdings in a company, the market becomes thinner with fewer potential buyers. Therefore, the price of the stock is likely to fluctuate more when new information on the firm arrives.

This prediction is tested using exclusion announcements from the Norwegian Government Pension Fund Global (GPFG), managed by Norges Bank Investment Management (NBIM). The GPFG has a portfolio of $1.8 trillion and equity interests in over 9,000 publicly traded companies, making it the largest sovereign wealth fund globally. Since 2006, the Council on Ethics has recommended excluding 181 companies for ethical reasons. The Executive Board of Norges Bank has implemented these recommendations and created a list of excluded companies including nuclear weapons manufacturers, coal mining companies, and companies associated with serious human rights abuses.

Three factors support the use of NBIM as a "laboratory" for studying the volatility around divestment. The first factor is governance. Exclusion decisions are made independently by the Council on Ethics using publicly available standards and guidelines, rather than by a portfolio manager based on current market information, thus providing exogenous timing for the announcements. The second factor is influence. Due to the high visibility of GPFG, many investors mirror NBIM's exclusions and adopt its investment principles. According to Rudolf and Yuan (2025), each NBIM exclusion generates about five additional institutional exclusions. Thus, the actions taken by NBIM serve as a leading indicator for broader withdrawals. The third key factor is the distinction between product and conduct exclusions. *Product-based* exclusions are determined by what a company manufactures, and not by a new event or crisis. As such, they provide a "clean" placebo for assessing only the effects of divestment. *Conduct-based* exclusions (i.e., human rights abuses, corruption, pollution) are accompanied by an ESG scandal. Therefore, any change in volatility after an announcement reflects both divestment and the scandal. If divestment raises volatility, then *product exclusions* should exhibit higher volatility, regardless of whether a scandal is present. Conversely, if only the scandal causes

*Corresponding author
simen.guttormsen@nmbu.no (S. Guttormsen)
ORCID(s): 0009-0002-2533-8561 (S. Guttormsen)

volatility changes, then no differences would be observed in post-exclusion volatility between product and conduct exclusions.

The data contradict these expectations. For all excluded firms, the variance ratio surrounding the announcement date is essentially equal to 1 (no change in volatility). In fact, more than two-thirds of excluded firms exhibit *less* volatility subsequent to exclusion. This conclusion is reinforced when comparing the two types of exclusions. As noted previously, *product exclusions* (the “scandal-free” control group) demonstrate no significant increase in volatility after exclusion. Similarly, while there appears to be some slight difference in variance ratios between conduct exclusions and the overall sample mean, this difference is statistically insignificant.

A difference-in-differences (DiD) estimate comparing the volatility of excluded firms vs. matched non-excluded peers with matching on GICS industry and size is similar for the two groups. An event-time DiD model estimated over both firm and event-time fixed effects has a relatively small magnitude of treatment effect and is statistically insignificant at the 5% level. Anticipation tests find no consistent trend toward increasing volatility before the announcement date of the inclusion/exclusion decision throughout the entire period under examination for product exclusions. This strengthens the original finding based solely on the timing of announcements to include the process by which a company becomes included or excluded. A series of equivalence tests shows that the true average variance ratio of returns falls within 30% of one for both the overall sample and product exclusions after controlling for cross-sectional dependence via a cluster bootstrap method. Power analysis indicates that the sample has sufficient power to detect economically significant changes in variance at conventional thresholds.

This paper contributes to three areas of research. First, it complements Berk and Van Binsbergen (2025), who show that divestment does not affect firms' cost of capital. The analysis here shows that divestment also does not raise firms' realized stock price volatility—the channel operating through thinner equity markets and decreased liquidity. Together, these findings indicate that both financial mechanisms predicted by the reduced-investor-base model—the cost-of-capital channel and the volatility channel—are empirically inactive at current levels of institutional divestment. Second, the method used to identify conduct/product exclusions (and thereby provide a controversy-free placebo test) is not found in previous studies of divestiture. Additionally, this paper uses a control group that matches the treatment group and a panel specification with time/firm fixed effects. Third, this study addresses a contemporary public policy issue. In November 2025, Norway's Parliament voted to put the GPFG ethics exclusion framework on hold, under international political pressure, with the pause to be reviewed in October 2026.

The rest of the paper is organized as follows. Section 2 describes the process that the GPFG uses for ethically excluding companies. Section 3 provides an overview of the theoretical and empirical literature on divestiture and volatility. Section 4 presents the data and methods. Section 5 reports the main findings, anticipation testing, matched control comparisons, panel DiD analyses, robustness testing, and equivalence analysis. Section 6 discusses the economic implications and limitations of the study. Section 7 concludes.

## 2. Institutional Background

### 2.1. GPFG and its ethical exclusion framework

Norway's Government Pension Fund Global (GPFG) is a sovereign wealth fund established using funds generated from oil revenues. At year-end 2025, GPFG had approximately $1.8 trillion under management, approximately 70% invested in equities of over 9,000 companies in 70 countries. NBIM, an investment unit of Norges Bank (the Central Bank of Norway), manages all investments for GPFG.

Ethical exclusion occurs through the Council on Ethics (*Etikkrådet*), which is an independent entity that determines if a company violates the Guidelines for Observation and Exclusion. The Council reviews each potential violation, then investigates further (typically taking 12–24 months) and recommends either observation or exclusion of the offending company. Based on these recommendations, the Executive Board of Norges Bank ultimately decides whether a company will be excluded from the fund. For conduct-based exclusions where there was sufficient data to determine both dates, the median time between the Council's recommendation to exclude and the Executive Board's public announcement was 171 days (mean: 204 days), with a range of 0 to 709 days.[1] Exclusion criteria fall into two broad categories:

[1] Data used for this calculation came from tracing the recommendations made by the Council on Ethics found at `www.etikkradet.no`. Recommendations regarding coal exclusions did not require review by the Council prior to the application of the coal revenue threshold. All other product criteria (nuclear weapons, tobacco products, cannabis products and cluster munitions) were reviewed by the Council.

- **Product-based criteria**: Production of specific weapons (nuclear, cluster munitions, anti-personnel mines), tobacco, and coal mining or coal-based power generation above revenue thresholds.
- **Conduct-based criteria**: Serious or systematic human rights violations, severe environmental damage, gross corruption, and other particularly serious violations of fundamental ethical norms.

The central feature of the identification strategy is the distinction between product and conduct exclusions. Product exclusions are based solely on the types of products that the corporation produces and not on some scandal or other controversy. In addition, the announcement itself does not contain any additional negative information about the company. For example, markets had long known that Lockheed Martin produces nuclear weapons before NBIM's exclusion decision.

Conduct exclusions, on the other hand, generally occur as a direct result of an ESG controversy (oil spills, investigations into corrupt practices, etc.) which may independently affect a firm's overall risk profile.

### 2.2. The exclusion sample

Over the time frame of 2006 through 2025, NBIM has excluded 181 firms—101 product exclusions and 80 conduct exclusions. Conduct exclusions and virtually all product exclusions follow recommendations from the Council on Ethics. Coal exclusions represent a rare exception where NBIM makes the decision directly. There is also significant unevenness in how the exclusions are distributed over time. 38 coal-related exclusions were announced on one date alone (April 14, 2016), immediately after a policy change to exclude companies generating revenue greater than 30% from thermal coal. Thus, there exists a natural clustering test. In addition, twenty of the excluded companies have since had their exclusions revoked, creating a natural symmetry test: if exclusion causes volatility, then revocation should cause reversals.

NBIM's usual investment in firms it excludes varies between 0.5% and 1.5% of issued capital. At first glance, such percentages may appear relatively low; nonetheless, Rudolf and Yuan (2025) demonstrate that excluding a company by NBIM results in four to five institutional investors also deciding to exclude the same company, amplifying the ownership withdrawal by a factor much larger than from NBIM alone. Consequently, even with this cascading effect, the overall divesting of the fund still represents only a small percentage of the firm's investor base.

### 2.3. The November 2025 Ethics Pause

In November 2025, Norway's Parliament (Storting) voted to put the framework for ethical exclusions temporarily on hold following diplomatic pressure from the United States triggered by NBIM's exclusion of Caterpillar Inc. on human rights grounds associated with the Israeli-Palestinian conflict. As a result, the ethical exclusion framework is currently being reviewed by a Parliamentary Committee. A report is expected in October 2026. Therefore, evidence of how these actions impact firm-level financial outcomes—or do not—is particularly relevant during this time.

## 3. Literature

Divestment may increase firm-level volatility due to various theoretical foundations. Merton et al. (1987) explains how investor recognition affects asset price equilibrium. When there are fewer investors for a given security, those who remain will take on more idiosyncratic risk. In turn, these investors demand higher expected returns. The effect increases as idiosyncratic volatility increases. Heinkel et al. (2001) apply Merton's theory to ethical investments, demonstrating that when "green" investors avoid companies that pollute, the remaining investors become the new pool of potential buyers. As such, the cost of capital rises, though their calibration demonstrates that approximately 25% of all investors would have to disinvest to affect company behavior substantially. Most recently, Pástor et al. (2021) develop an equilibrium-based model of sustainable investment. They demonstrate that green stocks receive lower expected returns because investors motivated by ESG values derive non-monetary benefits from owning shares. Pedersen et al. (2021) demonstrate that the equilibrium depends on the proportion of each type of investor in the marketplace, thereby producing ambiguous expectations regarding whether the ESG return premium exists. Edmans et al. (2022) determine the condition necessary for disinvestment to be socially desirable. Specifically, they show that disinvestment could actually make matters worse for companies facing less pressure to reform than previously existed. The models discussed here produce predictions mostly about expected returns and the cost of capital, not about realized volatility. Realized return volatility is produced by a second-step process where the reduced number of shareholders participating in the secondary market creates larger bid/ask spreads and larger price impacts. Therefore, a microstructure channel is created that increases realized return volatility regardless of any changes made to a firm's underlying fundamental cash flow risk. However, this microstructure channel operates through the reduced participation mechanism and is therefore independent of Merton et al.'s prediction concerning equilibrium pricing. Section 5 investigates the idiosyncratic channel directly through market-model residuals. A second theoretical branch examines divestment and its connection to *systematic* risk. Luo and Balvers (2017) demonstrate that social screens used together create a priced boycott-risk factor, indicating that divested companies carry a systematic risk premium unrelated to the idiosyncratic channel described above. Ansar et al. (2013) also argue that divestment campaigns stigmatize target firms, which leads to uncertainty related to future regulatory and reputational risk. Bolton and Kacperczyk (2021) provide evidence of a carbon premium in the cross-section of stock returns, consistent with investors requiring compensation for transition risk.

In contrast to expectations, research shows that divestment has minimal financial impact. Teoh et al. (1999) researched the anti-apartheid boycott of South African firms and found "no discernible effect" on the stock price of firms being boycotted. Hong and Kacperczyk (2009) showed that norm-constrained institutions avoided "sin stocks," creating an ongoing return premium; however, Blitz and Fabozzi (2017) indicated that the sin premium was small and statistically unstable after applying modern factor models. Chava (2014) found that companies whose environmental profiles raised concerns experienced a greater cost of equity, although it remained unclear how much of this reflected environmental risks themselves versus investor exclusion. Berk and Van Binsbergen (2025) provided the most current and rigorous analysis related to sustainable investing, demonstrating that sustainable investing has a negligible impact on a company's cost of capital. Even if socially responsible investors chose to divest from a particular company, the remaining investors could purchase those same shares at a trivially small discount since individual companies represented only a minuscule portion of the overall market portfolio. Hartzmark and Sussman (2019) demonstrated that mutual fund flows responded sharply to fund-level sustainability ratings, though high-sustainability funds did not outperform low-sustainability funds, suggesting investor demand rather than return premia drove the flow response. Hartzmark and Shue (2022) took their argument one step further, indicating that divesting from brown firms may actually have negative consequences due to green firms having lower impact elasticity. Nguyen et al. (2020) also identified that it was long-term institutional investors driving the relationship between CSR (Corporate Social Responsibility) and shareholder value. Therefore, they determined that the time horizon of an investor is more relevant than simply whether to include or exclude a specific company in an investment portfolio. Dimson et al. (2015), through their documentation of active ownership producing tangible behavioral changes in corporations, indicate that engagement may be a more effective means for influencing corporations than excluding them from portfolios.

The previous studies mainly focused on return effects of NBIM exclusions. For example, Atta-Darkua (2022) estimates the negative short-term announcement effects for the GPFG exclusion dates, with an average CAR of roughly −1.7%. On the other hand, Hoepner and Schopohl (2018) examined the exclusions from the GPFG and Swedish AP-funds. They found no significant return difference over longer horizons compared to benchmark portfolios. Berle et al. (2025) conducted the most extensive analysis of NBIM exclusions to date. They found that excluded firms generated 5–7% annual alpha relative to a Fama-French five factor model. This result is consistent with the "sin stock" premium literature. Their event study showed a negative CAR on the announcement day that reversed within a week. There was no permanent price effect of the exclusion announcements. In conclusion, their findings suggested that the exclusion announcements contained little additional information beyond what was already publicly available. Additionally, they showed that only 14% of excluded firms changed their operations sufficiently to have the exclusion revoked. Rudolf and Yuan (2025) documented a strong cascading behavior: each NBIM exclusion triggered about five additional institutional exclusions. However, they reported very small cumulative abnormal returns (−0.5%) over a three-day window. Therefore, this study complements Berle et al. (2025)'s work by examining the second moment (firm-level volatility). It investigates whether the reduction of an investor base leads to observable realized-volatility effects via market thinning and decreased liquidity—the microstructure channel implied by, but not modeled in, Merton et al. (1987).

There is a separate body of research showing that an ESG controversy has a direct link to firm-level risk. Xu et al. (2025) find evidence that firms' ESG performance reduces their stock price volatility and extreme risk, which shows that firms with poorly performing ESG ratings are inherently riskier. Therefore, this provides the alternative explanation central to the identification strategy: if NBIM's conduct-based exclusions arise from ESG controversies that independently increase volatility, then the measured effect could be attributed to the controversy instead of the divestment. In order to measure both of the above-mentioned effects—i.e., how much of the effect is due to the

**Table 1**
Exclusion Sample by Criterion

| Category | Criterion | Total | GARCH sample |
|---|---|---|---|
| *Product-based* | | | |
| | Coal | 63 | 52 |
| | Tobacco | 18 | 16 |
| | Nuclear weapons | 14 | 13 |
| | Cannabis | 4 | 4 |
| | Cluster munitions | 2 | 1 |
| | *Subtotal* | 101 | 86 |
| *Conduct-based* | | | |
| | Environment | 28 | 15 |
| | Human rights/conflict | 21 | 6 |
| | Human rights | 13 | 6 |
| | Corruption | 5 | 2 |
| | Other conduct | 13 | 10 |
| | *Subtotal* | 80 | 39 |
| **Total** | | **181** | **125** |

*Notes*: "Total" is the number of exclusion events compiled from NBIM's public database, 2006–2025. "GARCH sample" is the subset with sufficient daily return data for GARCH(1,1) estimation (minimum 150 estimation-window observations and 10 event-window observations). The drop from 181 to 125 reflects firms with thin trading histories, recently listed companies, or GARCH non-convergence.

divestment and how much is due to the controversy—the product/conduct decomposition is used in the empirical design.

# 4. Data and Methodology

## 4.1. Sample construction

This study uses an exclusion sample of 181 exclusion events that were publicly announced by NBIM between January 2006 and August 2025 and reported in its public exclusion database located on the website `nbim.no`. Each event is assigned to one of two types of exclusion, product or conduct, and then further to specific criteria such as coal, tobacco, nuclear weapons, human rights violations, environmental damage, and corruption. In addition, 20 reinstatement events are gathered for symmetry tests.

Daily price history is obtained for all excluded firms (both those excluded for their conduct and those excluded for their products), 92 matching control firms and the MSCI World ETF (URTH) for a market comparison. Prices are adjusted for both splits and dividends, and log returns are computed using the equation $r_t = \ln(P_t/P_{t-1})$.

For each exclusion event, pricing information is obtained for each firm from 252 trading days before the event until 60 days after the event. This ensures there will be enough historical price information available to estimate a GARCH model.

In total, the dataset has over 490,000 individual daily pricing observations of excluded firms and their matching control firms. Out of all exclusion events, GARCH estimation successfully converges for 125 (86 product and 39 conduct) events. The other 56 exclusion events are omitted because these events involve firms with too few pricing observations for successful GARCH estimation (for example, extremely low trading volumes or recently listed stocks), or because the GARCH estimation fails to converge when solving for the model parameters. The sample includes 30 different countries and 14 GICS industry groups.

Table 1 summarizes the exclusion sample by criterion and category.

### 4.2. Matched control group

For each firm excluded by NBIM, Refinitiv Eikon Screener is used to match an appropriate "control" firm. Control firms are selected based on the following criteria: (i) operate in the same four digit GICS Industry Group as the firm excluded from the sample, (ii) have a market capitalization at the time of the exclusion event that is within a factor of three of the excluded firm's market capitalization, and (iii) are listed in the same country if possible. If there are multiple candidate controls, the candidate with the smallest log market capitalization difference relative to the excluded firm is chosen, with a preference given to selecting a control firm from the same country. Of these, 106 yield enough price information about both firms in a pair (the excluded firm and its corresponding control firm) to qualify for the final dataset, while the remaining 75 involve firms operating in very specialized industries or in emerging markets where it would have been difficult to identify a suitable same-industry control firm. The 106 matched pairs use 92 unique control firms (median log market capitalization distance: 0.08; 92% same-country matches), substantially reducing the control-reuse and cross-country matching concerns that arise with manual pairing.

### 4.3. GARCH event study

The event-study methodology of MacKinlay (1997) is applied with an adaptation of the methodology for event-driven variance in Boehmer et al. (1991) and Savickas (2003). For every exclusion event $i$ the estimation window comprises all trading days within the time frame $[-272, -11]$, i.e., 252 observation periods from the event date, while the event window includes all trading days in the time frame $[-10, +20]$; this includes 31 observation periods. There is at least a ten-day separation between the two windows to prevent any potential pre-announcement effect contamination into the estimated baseline variance.

A GARCH(1,1)-model (Engle, 1982; Bollerslev, 1986) is fit for each firm over the estimation window:

$$r_t = \mu + \varepsilon_t, \quad \varepsilon_t = \sigma_t z_t, \quad z_t \sim N(0,1) \tag{1}$$

$$\sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2 \tag{2}$$

The primary function of fitting the GARCH model is as a quality filter for the return series data. Convergence implies there exists sufficient variation in the return series to allow the conditional variance model to accurately determine its parameters. The variance ratio tests, described in Section 4.5, rely upon simple estimates of sample variance derived from returns rather than fitted values of conditional variance via the GARCH model. As such, the results are not dependent upon the GARCH specification. Those events failing to achieve GARCH convergence are removed because the underlying return series is likely either too sparse or too irregular to reliably estimate variance through any method.

### 4.4. Idiosyncratic variance ratio

As an additional complement to the specification that tests for idiosyncratic risk as identified in prior theoretical work, variance ratios on residual returns produced by the market model are used to identify which firms exhibit increased volatility due to firm-specific risk factors rather than common risk factors. The simple market model is estimated on the estimation window:

$$r_{i,t} = \alpha_i + \beta_i r_{m,t} + \varepsilon_{i,t} \tag{3}$$

where $r_{m,t}$ is the S&P 500 return. Residuals $\hat{\varepsilon}_{i,t}$ are estimated over both the event window and the estimation window using the coefficients from the estimation window. Variance ratios estimated on residual returns ($\mathrm{VR}_i^{\mathrm{idio}} = \hat{\sigma}^2_{\varepsilon,\mathrm{event}}/\hat{\sigma}^2_{\varepsilon,\mathrm{est}}$) represent the idiosyncratic volatility experienced during the event period relative to that experienced during the estimation period.

### 4.5. Variance ratio test

The principal test statistic is the variance ratio (VR):

$$\mathrm{VR}_i = \frac{\hat{\sigma}^2_{\mathrm{event},i}}{\hat{\sigma}^2_{\mathrm{est},i}} \tag{4}$$

where $\hat{\sigma}^2_{\mathrm{event},i}$ is the realized variance in the event window and $\hat{\sigma}^2_{\mathrm{est},i}$ is the realized variance in the estimation window. Under the null of no event-induced volatility, $\mathrm{VR}_i \approx 1$. Elevated event-window volatility yields $\mathrm{VR} > 1$.

**Table 2**
Variance Ratio Tests Around NBIM Exclusion Announcements

| | $N$ | Mean VR | Median VR | Std. Dev. | $t$-stat | $p$-value |
|---|---|---|---|---|---|---|
| **Full sample** | 125 | 1.038 | 0.772 | 1.057 | 0.403 | 0.687 |
| *By exclusion type:* | | | | | | |
| Product-based | 86 | 0.874 | 0.745 | 0.659 | −1.773 | 0.080 |
| Conduct-based | 39 | 1.400 | 0.933 | 1.576 | 1.585 | 0.121 |

*Notes*: Variance ratio (VR) is the ratio of realized variance in the event window $[-10, +20]$ to realized variance in the estimation window $[-272, -11]$. Under the null of no event-induced volatility, VR $= 1$. The $t$-statistic tests $H_0$: $E[\text{VR}] = 1$.

The following cross-sectional variance ratio test is used according to Białkowski et al. (2008):

$$t_{\text{VR}} = \frac{\overline{\text{VR}} - 1}{s_{\text{VR}}/\sqrt{N}} \tag{5}$$

where $\overline{\text{VR}}$ is the average cross-sectional variance ratio and $s_{\text{VR}}$ is the cross-sectional standard deviation of individual VRs. The above test assumes no correlation or dependency across the sample of variance ratios. Because firms excluded on the same date share common shocks, cross-sectional dependence is addressed separately through a cluster bootstrap that resamples announcement dates with replacement (Section 5.7).

### 4.6. Difference-in-differences

For every matched pair of firms $(i, j)$ where $i$ is the excluded firm and $j$ is the control:

$$\text{DiD}_{\text{VR}} = \text{VR}_i^{\text{excluded}} - \text{VR}_j^{\text{control}} \tag{6}$$

In this case, the same event date (the date of the exclusion announcement) is used to estimate both variance ratios. Under the null hypothesis that divestment adds nothing beyond the underlying controversy, there should be no difference in the variance ratios ($E[\text{DiD}_{\text{VR}}] = 0$).

To support the DiD for cross-sectionally-matched pairs, a panel approach is also used. For each of the excluded-control pairs, 20-day rolling realized variance is calculated based on daily returns. Both series are aligned on event-time. The panel regression is:

$$\log(\text{RV}_{i,t}) = \alpha_i + \gamma_t + \beta \cdot (\text{Excluded}_i \times \text{Post}_t) + \varepsilon_{i,t} \tag{7}$$

This includes firm-specific fixed effects, denoted as $\alpha_i$, and event-time fixed effects, denoted as $\gamma_t$. $\text{Post}_t$ is equal to 1 if $t$ falls within the range $[0, 20]$ and 0 otherwise. A pre-treatment period that extends to $t = -60$ is used. Clustered standard errors are used at the pair level to account for shared controls. Additionally, replacing the post-period indicator with event-time dummies traces out the dynamic treatment effect.

## 5. Results

### 5.1. Full sample: no aggregate volatility effect

Results for the median and average variance ratios are shown in Table 2. Using the full sample of 125 exclusion events, the median variance ratio (VR) is 0.77. More than two-thirds (68%) of firms show lower variance in their stock prices after an announcement about exclusion compared to before. The average VR of 1.04 is heavily influenced by several extreme right-tail outlier values but is statistically indistinguishable from 1 ($t = 0.40$, $p = 0.69$). The null hypothesis that there is no effect on volatility due to the firm's exclusion announcement therefore cannot be rejected at any conventional significance level.

**Table 3**
Product Variance Ratios by Criterion

| Criterion | $N$ | Mean VR | Median VR | Std. Dev. | $t$-stat | $p$-value |
|---|---|---|---|---|---|---|
| Coal | 52 | 0.776 | 0.622 | 0.432 | −3.748 | $< 0.001$ |
| Tobacco | 16 | 0.796 | 0.772 | 0.413 | −1.974 | 0.067 |
| Nuclear weapons | 13 | 1.071 | 0.931 | 0.869 | 0.296 | 0.772 |
| Cannabis | 4 | 0.873 | 0.835 | 0.147 | −1.724 | 0.183 |
| Cluster munitions | 1 | 4.688 | — | — | — | — |
| *Excl. coal batch* | 53 | 0.904 | 0.747 | 0.759 | −0.921 | 0.362 |
| *Non-coal product* | 34 | 1.025 | 0.812 | 0.867 | 0.163 | 0.871 |

*Notes*: The coal batch refers to 33 firms excluded on April 14, 2016 following adoption of the 30% coal-revenue threshold. "Excl. coal batch" removes these 33 firms from the product sample. "Non-coal product" includes only tobacco, nuclear weapons, cannabis, and cluster munitions.

## 5.2. Product vs. conduct: identification through exclusion type

The cleanest sample of all (in terms of identifying how firms respond) is the 86 "product" exclusions—those with no scandalous news prior to announcement which could have potentially interfered with the divestiture process. In such instances the average VR was 0.87—below one and just marginally statistically different from 1 at the 10% level using t-tests ($t = -1.77$, $p = 0.08$). This is precisely the opposite of what was expected; theoretically, product exclusions were thought to produce higher post-announcement volatility, whereas they produced less volatility instead. The equivalence test in Section 5.8 verifies that the average VR of the product sample falls inside of a 30% interval around unity ($p = 0.047$ when the same-date dependence is accounted for via cluster bootstrapping). The power calculations also indicated that the product sample had an 80% chance of detecting a 21% increase in variance.

The 39 conduct exclusions are less helpful in terms of establishing evidence. 54% of the conduct-excluded companies experience decreased volatility after their exclusion from NBIM's portfolio. The median VR is 0.93, below 1, but a small number of extremely large outlier values (e.g., Adani Ports, VR = 7.04; China State Construction, VR = 7.75) are enough to pull the mean above 1 (mean = 1.40). The $t$-statistic for these 39 observations is only 1.59 ($p = 0.12$), and because there is such high variability (SD = 1.58), the conduct-excluded sample cannot establish equivalence even at an equivalence margin as wide as $\delta = 0.5$. Moreover, achieving 80% power would require a true VR of at least 1.73. This subsample is therefore suggestive—the pattern is consistent with the controversy driving the noise—but it cannot carry much weight on its own.

Figure 1 illustrates this contrast through rolling variance ratios.

*Sub-product splits.* Table 3 reports variance ratios by product criterion. Coal exclusions ($N = 52$, mean VR = 0.78, $p < 0.001$) and tobacco exclusions ($N = 16$, mean VR = 0.80, $p = 0.07$) fall below unity, while nuclear weapon exclusions ($N = 13$, mean VR = 1.07, $p = 0.77$) are near unity. The single cluster munitions observation (VR = 4.69, $N = 1$) is uninformative. The below-unity coal VRs may be mechanical: if the estimation window $[-272, -11]$ contains elevated volatility from impending coal-policy debates or commodity shocks, the denominator will be inflated and the ratio will fall below one. Three points alleviate this concern. First, the matched-control DiD (Section 5) accounts for sector-by-date volatility shocks and shows no effect for product exclusions. Second, when the April 2016 coal batch is excluded (33 firms excluded on a single day after the policy threshold was met), the product mean VR equals 0.90 ($p = 0.36$): the null holds either way. Third, non-coal product exclusions ($N = 34$, VR = 1.02, $p = 0.87$) show no effect; the null does not rely on coal.

*Idiosyncratic variance ratio.* To identify the channel through which firm-specific volatility affects stock prices as predicted by theory, variance ratios are calculated using residuals from the market model (as described in Section 4). The S&P 500 (SPY) is used as the market proxy to estimate $r_{i,t} = \hat{\alpha}_i + \hat{\beta}_i r_{m,t} + \hat{\varepsilon}_{i,t}$ over each estimation window. All 125 firms with successful GARCH estimation generate idiosyncratic VRs.[2] The quantitative results from the idiosyncratic

[2] The S&P 500 (SPY) is used as the market proxy for the idiosyncratic specification since its history includes all 125 events from 2006 forward. The MSCI World ETF (URTH) starts in 2012 and thus excludes 19 of the earliest events.

**Table 4**
Pre-Event Variance Ratios: Testing for Anticipation

| Window | Category | $N$ | Mean VR | Median VR | $t$-stat | $p$-value |
|---|---|---|---|---|---|---|
| Main $[-10, +20]$ | All | 125 | 1.038 | 0.772 | 0.403 | 0.687 |
| | Product | 86 | 0.874 | 0.745 | −1.773 | 0.080 |
| | Conduct | 39 | 1.400 | 0.933 | 1.585 | 0.121 |
| Pre $[-60, -10]$ | All | 127 | 2.289 | 0.917 | 2.400 | 0.018 |
| | Product | 86 | 1.193 | 0.935 | 1.553 | 0.124 |
| | Conduct | 41 | 4.588 | 0.799 | 2.247 | 0.030 |
| Pre $[-30, -10]$ | All | 127 | 0.960 | 0.717 | −0.542 | 0.588 |
| | Product | 86 | 0.934 | 0.700 | −0.698 | 0.487 |
| | Conduct | 41 | 1.013 | 0.752 | 0.113 | 0.911 |

*Notes*: Each row reports the cross-sectional mean and median variance ratio for the indicated window. The estimation window is 252 trading days ending immediately before the pre-event window. A pre $[-120, -10]$ window is not reported because most events lack sufficient data, but yields qualitatively similar results for the subset that does.

specification are very similar to those of the raw specifications: product idiosyncratic VR is 0.89 (raw: 0.87), conduct is 1.41 (raw: 1.40), and the full sample is 1.05 (raw: 1.04). The mean market-beta for firms excluded is 0.56 with a typical $R^2$ of approximately 0.11; hence these firms have moderate but non-negligible market exposure. Eliminating this common variability does not affect the conclusions; therefore, idiosyncratic volatility remains unaffected by the announcement of exclusions.

## 5.3. Anticipation and full-timeline evidence

NBIM's exclusion procedure usually takes 12–24 months from when an investigation is initiated to public announcement. For conduct-based exclusions, the median time between the Council on Ethics recommendation and the Board's announcement is 171 days (see Section 2). If information about a planned exclusion leaks before the announcement date, volatility in the excluded firm's stock may have already increased, biasing the main test against detecting any statistically significant effect. To mitigate this risk, three alternative tests are presented.

Variance ratios are calculated in pre-event windows: $[-120, -10]$, $[-60, -10]$, and $[-30, -10]$ trading days from the event. These are reported in Table 4. Product exclusions do not have increased variance ratios in any of the pre-event windows ($p > 0.12$ in all cases). Conduct exclusions show a significant mean VR of 4.59 ($p = 0.03$) in the $[-60, -10]$ window; however, the median VR is 0.80, indicating that there are a small number of extreme values, consistent with an ongoing ESG scandal and not leakage of the NBIM decision. The $[-30, -10]$ window shows no significance for either category.

Exclusion types are also categorized based on their potential for causing surprise. The batch announcement type of exclusions (exclusion of more than three companies simultaneously, $N = 91$) is likely to be less surprising than single company exclusion types ($N = 34$) since in most cases the batch decision follows a policy change (i.e., the coal threshold from 2016). Neither group has statistically significant variance ratios (batch: $p = 0.73$; single exclusion: $p = 0.32$). Similarly, no statistical difference is found in either first-ever use or routine application of a criterion ($N = 58$ and $N = 67$ respectively; first-ever: $p = 0.11$; routine: $p = 0.33$).

In addition, Figure 2 presents the median of the rolling 20-day VRs over time in the cross-section from 180 days before to 60 days after the announcement. The product panel has a flat trajectory with no pre-announcement trend, indicating that the results are not due to "information leakage". Although the conduct panel has greater variability, it is likely related to the timing of the scandal itself and not to the fact that the firm was excluded.

*Long-term test.* In order to examine if the exclusion process as a whole (and not simply its announcement) has had an effect on volatility, realized six-month volatility prior to the first announced signal of the exclusion (the Council on Ethics recommendation for conduct exclusions or the initial announcement for product exclusions) is compared to realized six-month volatility after the announcement. Of the 125 firms with successful GARCH estimation, 119

**Table 5**
Difference-in-Differences: Excluded vs. Matched Control Firms

| | $N$ | Mean | Median | Winsorized | $t$-stat | Wilcoxon $p$ |
|---|---|---|---|---|---|---|
| **All pairs** | 79 | +0.145 | +0.058 | +0.072 | 1.27 | 0.361 |
| Product | 49 | −0.011 | +0.026 | −0.020 | −0.12 | 0.867 |
| Conduct | 30 | +0.400 | +0.140 | +0.369 | 1.59 | 0.213 |

*Notes*: DiD(VR) = $\text{VR}_{\text{excluded}} - \text{VR}_{\text{control}}$. Each excluded firm is matched to a same-GICS-industry-group, size-matched control firm not excluded by NBIM (92 unique controls; 92% same-country). Both VRs are computed around the same event date. Winsorized mean clips at the 5th and 95th percentiles. Under $H_0$: $E[\text{DiD}] = 0$.

have adequate return histories for this analysis. The realized six-month volatility ratio for product exclusions is 1.000 ($t = 0.030$, $p = 0.98$); realized volatility is essentially unchanged before and after the announcement. Realized six-month volatility ratios for conduct exclusions are slightly higher at 1.150 ($p = 0.11$), consistent with being generated by the ongoing controversy surrounding the firm and not by the exclusion.

Together, these analyses expand the previously identified null result from the announcement date to include all of the firms' timelines during which they were subject to an exclusion process. For product exclusions, neither the 12–24 month time frame for each type of investigation nor the announcement date itself produced statistically significant increases in realized volatility. Therefore, the null cannot be attributed to information leakage: the divestment channel is inactive across the entire exclusion process.

## 5.4. Matched controls: no detectable divestment effect beyond controversy

Table 5 reports the results from the DiD method using the 79 matched pairs with enough information such that each of the excluded companies has a same GICS industry, size-matched control company that was also not excluded by NBIM (92 unique controls; 92% same-country). The median DiD variance ratio in Table 5 is +0.058—a close-to-zero value. Also, the Wilcoxon signed-rank test does not reject the null hypothesis ($p = 0.36$). The DiD mean for the product exclusion category (49 pairs) is −0.01 ($t = -0.12$, $p = 0.91$), which indicates essentially zero effect. Conduct exclusions yield a noisy positive mean of +0.40; however, no statistical significance is found for this estimate ($p = 0.12$). Winsorizing at the 5th/95th percentiles barely reduces the magnitude of this mean ($p = 0.09$). Therefore, excluded firms experience volatility comparable to their matched controls, indicating that there is no detectable volatility effect added by the divestiture process beyond that caused by the original controversy. Figure 3 illustrates that the scatter plots of excluded vs. control VRs cluster closely about the 45-degree line.

## 5.5. Panel event-time difference-in-differences

DiD comparing individual variance ratios across matched pairs is less efficient than using the full panel of daily observations. Thus, for each exclusion-control pair, 20-day rolling realized variance is calculated from daily returns; then, the entire panel of daily observations is used to perform a regression as shown by Equation (7), including firm and event-time fixed effects. The pre-exclusion period extends up until $t = -60$ days. $\beta$ measures how much more the log of realized variance has increased for the excluded firm relative to the control firms over the event window $[0, +20]$. Standard errors are clustered at the pair level to account for shared controls.

Table 6 shows the findings. For the full sample (105 pairs, 15,068 firm-day observations) $\hat{\beta} = 0.138$ (SE = 0.082, $p = 0.09$). For product exclusions (67 pairs) $\hat{\beta} = 0.177$ ($p = 0.10$). For conduct exclusions (38 pairs) $\hat{\beta} = 0.085$ ($p = 0.52$). All three estimates are insignificantly different from zero at the 5% significance level. However, both the full-sample and product samples yield marginally significant evidence at the 10% significance level.

The TOST equivalence test cannot confirm an increase of 20% or greater in realized variance for product exclusions ($p_{\text{TOST}} = 0.149$). Both tests show internal consistency: they indicate that no increase greater than about 20% can be rejected for product exclusions. The cluster bootstrap in Section 5.8 provides additional support for these findings: equivalence holds at $\delta = 0.3$ for both the full sample ($p = 0.015$) and product ($p = 0.047$); however, it does not hold at $\delta = 0.2$. Therefore, the product null is bounded: any potential effect is below 30%. The small discrepancy between the cross-sectional DiD test (product: $p = 0.91$) and the panel specification (product: $p = 0.10$) may stem from

**Table 6**
Panel Event-Time Difference-in-Differences

| Sample | Pairs | $N_{\text{obs}}$ | $\hat{\beta}$ | SE | $p$-value |
|---|---|---|---|---|---|
| All | 105 | 15,068 | +0.138 | 0.082 | 0.093 |
| Product | 67 | 9,576 | +0.177 | 0.106 | 0.096 |
| Conduct | 38 | 5,492 | +0.085 | 0.133 | 0.523 |

*Notes*: Dependent variable is $\log(\text{RV}_{i,t})$ where RV is 20-day rolling realized variance. $\text{Post}_t = 1$ for event days $[0, +20]$; the pre-period extends to $t = -60$. All regressions include firm and event-time fixed effects. Standard errors clustered at the pair level to account for shared controls. Controls matched on GICS industry group, market capitalization (factor of 3), and country.

**Table 7**
Reinstatement (Revocation) Analysis

| Test | $N$ | Mean | $t$-stat | $p$-value |
|---|---|---|---|---|
| Reinstatement VR ($H_0$: $= 1$) | 16 | 1.031 | 0.247 | 0.809 |
| VR(excl.) − VR(reinst.) ($H_0$: $= 0$) | 13 | 0.092 | 0.317 | 0.757 |

*Notes*: The reinstatement VR test asks whether variance changes around reinstatement dates. The asymmetry test compares each firm's exclusion-period VR to its reinstatement-period VR. Under the null of no divestment effect, both statistics should be insignificant.

the construction of realized variance. This involves calculating a series of rolling-window return variances, which creates mechanical autocorrelation within each window of observation and increases the effective sample size. As such, the cross-sectional VR test—which uses a single non-overlapping variance ratio per firm—represents the most conservative primary test statistic.

Figure 4 provides an alternative estimation strategy for the DiD model by replacing $\text{Post}_t$ with weekly bin dummies in the pre-period and daily bin dummies near the event, which traces out the dynamic treatment effect across event time. The pre-announcement coefficients cluster around zero for both product and conduct subsamples, providing visual evidence for the parallel-trends assumption. No post-announcement spike is visible.

## 5.6. Reinstatement symmetry test

Portfolio readmissions (reinstatement events) also serve as an additional mechanism for testing. If Merton's risk-sharing channel exists, then exclusion should permanently raise equilibrium volatility, and therefore the effects of exclusion should be reversed by the reinstatement. Table 7 shows that there are 16 firms in which reinstatement has occurred. The average VR for reinstated firms is 1.03 ($t = 0.25$, $p = 0.81$); thus reinstatement, similar to exclusion, is a non-event from a volatility perspective. The difference between each firm's exclusion VR and its reinstatement VR averages 0.09 ($t = 0.32$, $p = 0.76$), consistent with neither of the two events affecting risk.

## 5.7. Robustness of the null

The issue raised by the positively skewed VR distribution is whether outliers could mask a genuine effect or generate a spurious one. This is addressed in Table 8, using distribution-free tests. For product exclusions, the Wilcoxon signed-rank test clearly rejects VR = 1 ($p < 0.001$); however, contrary to expectation, the median VR of 0.74 indicates lower post-announcement variance. The sign test also supports that fact: only 22 of the 86 product events have VR $> 1$ ($p < 0.001$). Winsorized mean values (at the 5th and 95th percentiles) and 10% trimmed mean values indicate the same thing. For conduct exclusions, all robust tests are nonsignificant ($p > 0.14$). Therefore, outlier sensitivity further strengthens the null hypothesis: the product results appear stable and directionally opposite to the theoretical prediction.

**Table 8**
Outlier-Robust Variance Ratio Tests

| | Wilcoxon | | Sign test | | Winsorized | |
|---|---|---|---|---|---|---|
| | $W$ | $p$ | $k/N$ | $p$ | Mean | $p$ |
| Product ($N = 86$) | 882 | $< 0.001$ | 22/86 | $< 0.001$ | 0.804 | $< 0.001$ |
| Conduct ($N = 39$) | 356 | 0.644 | 18/39 | 0.749 | 1.229 | 0.141 |
| All ($N = 125$) | 2657 | 0.002 | 40/125 | $< 0.001$ | 0.923 | 0.114 |

*Notes*: Wilcoxon signed-rank test and sign test both test $H_0$: median VR = 1. $k/N$ is the number of events with VR > 1. Winsorized mean clips at the 5th and 95th percentiles; its $p$-value is from a one-sample $t$-test against 1.

Null stability holds in all alternative event windows ($[-1, +1]$, $[-5, +5]$, $[-20, +40]$); in all alternative GARCH models (GJR-GARCH, Glosten et al., 1993; EGARCH, Nelson, 1991); in all pre/post-2016 subsamples; in all pre/post-COVID subsamples; and in a permutation test using 1,000 randomly generated pseudo-event dates. The empirical $p$-value is 0.73. Full details of each of these supplemental tests can be made available upon request.

*Cluster bootstrap inference.* One concern about the firm-level cross-sectional test is that firms excluded on the same day may have shared shocks or other influences that inflate the effective sample size. Therefore, a cluster bootstrap method is used to sample the announcement dates. This procedure retains the within-date dependence structure while retaining all firm-level observations. 10,000 replications of the cluster bootstrap are performed. The cluster bootstrap standard error is approximately 38% larger than the naive standard error for product and 23% larger for the full sample. Despite this increased variation due to clustering, the null result is unchanged: the cluster-bootstrap $p$-value for the full sample is 0.74 (naive $p = 0.69$) and for product it is 0.14 (naive $p = 0.08$). The cluster bootstrap 95% confidence interval for product is [0.74, 1.13], which excludes any plausible economic increase in VR. Finally, collapsing VRs into "waves" based on the announcement date (i.e., each wave represents one observation per day) yields 35 independent waves with a mean VR of 1.08 ($p = 0.47$). These results confirm the null hypothesis when an alternate dependence correction scheme is applied.

### 5.8. Equivalence tests and power analysis

Failure to reject $H_0$: VR = 1 is not sufficient in itself to prove that the true effect is zero. It can also reflect low statistical power. Therefore, in addition to the null-hypothesis tests, two one-sided tests (TOST) for equivalence and a formal power analysis are performed.

The TOST procedure tests whether the true mean VR lies in the interval $[1-\delta, 1+\delta]$ at significance level $\alpha = 0.05$. Table 9 reports results for three values of $\delta \in \{0.2, 0.3, 0.5\}$ using both naive (independence assumed) and cluster-bootstrap standard errors. For $\delta = 0.2$ the naive TOST establishes full-sample equivalence ($p = 0.044$); however, this result does not hold when the cluster-bootstrap method is used ($p = 0.086$), confirming that calendar clustering inflates naive precision. The TOST procedure with $\delta = 0.3$ shows equivalence for all cases: the cluster-bootstrap TOST rejects the non-equivalence null for both the whole data set ($p = 0.015$) and products ($p = 0.047$). The conduct subsample never achieves equivalence; its very high variability accounts for this, rather than a real effect.

Figure 5 shows 90% and 95% confidence intervals for the mean VR in each category based on naive standard errors. Under the cluster bootstrap, the 95% CI for the product subsample is [0.74, 1.13]—still excluding any economically meaningful volatility increase. The full-sample cluster-bootstrap CI is [0.86, 1.31], and the conduct CI of [1.01, 1.90] lies entirely above unity, but this should not be interpreted as evidence for a divestiture effect: the conduct subsample is small ($N = 39$), dominated by outliers tied to major scandals (Section 5), and confounded by the controversy that triggered exclusion. The elevated conduct mean reflects these confounds, not the institutional exit.

Figure 6 illustrates the power curves for the three categories—the probability of rejecting $H_0$: VR = 1 given various values of the true VR. With an $N = 86$ and $\sigma = 0.66$ in the product subsample, the test has 80% power to detect a 21% variance increase. The full sample detects a 27% increase. With $N = 39$ and $\sigma = 1.58$ in the conduct subsample, it would require a 73% variance increase to achieve 80% power. Thus, the conduct subsample's results are likely to be inconclusive.

**Table 9**
TOST Equivalence Tests: Naive and Cluster Bootstrap

| | $\delta = 0.2$ | | $\delta = 0.3$ | | $\delta = 0.5$ | |
|---|---|---|---|---|---|---|
| | $p_{\text{TOST}}$ | Equiv.? | $p_{\text{TOST}}$ | Equiv.? | $p_{\text{TOST}}$ | Equiv.? |
| *Naive (independence):* | | | | | | |
| All ($N = 125$) | 0.044 | Yes | 0.003 | Yes | $< 0.001$ | Yes |
| Product ($N = 86$) | 0.149 | No | 0.008 | Yes | $< 0.001$ | Yes |
| Conduct ($N = 39$) | 0.786 | No | 0.655 | No | 0.345 | No |
| *Cluster bootstrap (35 clusters):* | | | | | | |
| All ($N = 125$) | 0.086 | No | 0.015 | Yes | $< 0.001$ | Yes |
| Product ($N = 86$) | 0.230 | No | 0.047 | Yes | $< 0.001$ | Yes |
| Conduct ($N = 39$) | 0.803 | No | 0.666 | No | 0.334 | No |

*Notes*: Two one-sided tests (TOST) for equivalence. The null hypothesis is $|E[\text{VR}] - 1| \geq \delta$; rejection implies the true VR lies within $[1 - \delta, 1 + \delta]$. Test at $\alpha = 0.05$. "Naive" uses the cross-sectional standard error assuming firm-event independence. "Cluster bootstrap" resamples announcement dates with replacement (10,000 replications), preserving within-date dependence, and uses the bootstrap standard error for the TOST $t$-statistics. Product events span 18 unique dates; conduct events span 23 unique dates.

# 6. Discussion

## 6.1. Magnitudes and the Merton channel

The theory predicts that divesting will increase volatility when the exiting investors represent a substantial percentage of the company's investor base. Typically, Norges Bank Investment Management (NBIM) owns about 0.5–1.5% of its investments. Even using the 5× cascade described in Rudolf and Yuan (2025), the aggregate divestment represents roughly 3–8% of institutional ownership. Therefore, thousands of other investors—including other institutions and individual investors—can absorb the shares being divested without impacting the equilibrium of risk sharing among investors. This conclusion supports the intuition from Berk and Van Binsbergen (2025): in highly liquid, globally traded equities, there are many potential sources of capital available to invest in any specific equity compared to the quantity of capital that ESG-motivated investors can potentially withdraw. The study does not contradict Merton et al. (1987) as a theoretical concept; however, the study indicates that at the sizes of institutional ownership typically seen today, the Merton et al. (1987) mechanism produces no discernible changes in volatility due to ESG-motivated divestments.

## 6.2. The product/conduct decomposition as a placebo design

The "product/conduct" separation offers a relatively rare opportunity for a placebo test that few divestiture-related studies can offer. Most studies of divestment effectiveness suffer from endogeneity. Firms are excluded due to some adverse event (a scandal, a regulatory action), which impacts both price and risk independent of the investor's divestment decision. Separating product-based removals—controversy-free exclusions where the "news" is merely the characteristics of the product, well understood by financial markets for years—substantially reduces the potential confounds associated with divestiture studies, although some residual confounding remains since many coal exclusions were threshold-based policy decisions and therefore potentially partially anticipated. The fact that product exclusions produce null (or even below-unity) VRs, whereas conduct exclusions produce noisier, positive VRs, is evidence that the controversy and not the divestiture is likely responsible for any observed effects.

## 6.3. Implications for the ESG divestment debate

Together with Berk and Van Binsbergen (2025), these results close both financial channels through which theory predicts divestment imposes costs on firms that have been excluded from investment portfolios. The reduced-investor-base framework identifies two mechanisms: fewer holders should raise required returns (the first-moment, cost-of-capital channel), and fewer holders should thin the secondary market, increasing price variability (the second-moment, volatility channel). Berk and Van Binsbergen (2025) find no detectable effect on the first mechanism. This paper similarly finds no detectable effect on the second mechanism. At present, institutional ownership levels suggest the financial case for divestment as a means of imposing significant financial costs on excluded firms is relatively weak.

This does not imply that divestment is pointless. As noted previously, Rudolf and Yuan (2025) document the ability of divestment decisions made by large asset owners, such as Norway's sovereign wealth fund (GPFG), to influence other asset owners' decision-making processes. Therefore, the primary mechanism through which divestment appears to operate may be based upon political and reputational considerations rather than purely financial ones. The extent to which increased scale or coordination among multiple asset owners might trigger significant financial impacts on excluded firms remains uncertain.

### 6.4. Limitations

A few caveats qualify the interpretation of the null hypothesis. First, NBIM does not disclose when each trade is made. Therefore, while the announcement shock is captured using an event window of $[-10, +20]$, it is unclear whether this window allows enough time for all of the trading necessary to clear the market; this is also why two methods (the anticipation tests and the six-month horizon comparison) are used in Section 5.3 to establish the timing of the ownership changes. Second, with respect to the idiosyncratic variance ratio (VR), a single U.S. market factor (S&P 500) is used for a sample of 30 countries. This implies that there could be other country or currency-specific factors that contribute to the residuals of the VR model. However, the estimates of VR are very similar to those based on raw data, and therefore the single market proxy does not explain away the null hypothesis. Third, realized volatility is estimated rather than implied volatility because options on many of the excluded firms are illiquid. Fourth, the subsample of conduct exclusions is sufficiently small and fragmented to make statistical inference difficult, and therefore equivalence to the null cannot be shown even at $\delta = 0.5$. Finally, the results provide evidence on one institutional investor. It remains for future research to determine whether a coordinated withdrawal from equity markets by multiple large institutional investors would result in significantly different conclusions.

### 6.5. Policy implications: Norway's 2025 ethics review

The result that NBIM's exclusions have no impact on firm-level volatility has serious ramifications for the ongoing parliamentary review of the ethical exclusion framework. If the merit of the framework can be judged by its negative financial effects on the companies it excludes, then this argument is very weak. If instead the merit of the framework lies with its signaling and coordination roles—aligning institutional investors to common ethical standards, sparking a national dialogue about ethics and generating reputation-based risks for companies violating these ethical standards—then its merits must be determined based on those non-economic characteristics. The financial rationale for either maintaining or dismantling the framework appears equally thin.

## 7. Conclusion

Based on data from 181 NBIM exclusion events over a span of nearly 20 years, no statistically significant evidence is found of an increase in volatility due to the announcement of divestiture by the world's largest sovereign wealth fund or as a result of an investigation conducted prior to the decision to exclude. The product/conduct decomposition—the primary method for identifying the causal relationship—shows that the product-based, "controversy-free" placebo events show no impact, while the noisier "conduct-based" events reflect the controversy associated with the event and not the institutional exit itself. When a matched control analysis employing industry- and size-matched controls was used to produce an estimate of the DiD for product exclusions at the cross-sectional level, it produced estimates that were effectively zero. At the panel event-time level, with firm-specific fixed effects, while there was some evidence of economic significance, this did not reach statistical significance at the 5% level. Anticipation tests find no pre-announcement drift for product exclusions. Equivalence tests indicate that the true mean value for VR (variance ratio) lies within 30% of unity for both the full sample and for product exclusions. This finding holds regardless of whether a cluster bootstrap was employed to account for the cross-sectionally interdependent structure of same-date exclusion batches. The results are robust across GARCH specifications, event windows, subsamples, and a permutation test.

The results close the second of two financial mechanisms that economic theory predicts arise when firms are excluded from an investor's portfolio through divestiture. Berk and Van Binsbergen (2025) demonstrate there is no evidence that divestment increases the cost of capital; this study demonstrates similarly that there is no increase in realized volatility. Although theoretically the thinner markets and lower liquidity implied by (though not directly modeled in) Merton et al. (1987) can result from a reduced ownership base, current ownership magnitudes appear insufficient to activate either financial channel. While divestiture may offer other advantages—political, reputational, and coordinating—at NBIM's level of ownership, the financial case for imposing costs on excluded companies is not supported by the data.

## Data availability

The NBIM exclusion list and associated decision documents are publicly available from the Norwegian Government Pension Fund Global's ethical exclusion database at `nbim.no`. Council on Ethics recommendation dates are available at `etikkradet.no`. Daily equity prices and market-capitalization data were obtained from LSEG (Refinitiv Eikon) under licence. Restrictions apply to the availability of these data, which are available from the author with the permission of LSEG.

## Declaration of generative AI and AI-assisted technologies in the manuscript preparation process

During the preparation of this work the author used Claude Opus 4.6 (Anthropic) in order to proofread the manuscript, improve readability, and review code. After using this tool, the author reviewed and edited the content as needed and takes full responsibility for the content of the published article.

## References

Ansar, A., Caldecott, B., Tilbury, J., 2013. Stranded assets and the fossil fuel divestment campaign: what does divestment mean for the valuation of fossil fuel assets? Stranded assets and the fossil fuel divestment campaign .

Atta-Darkua, V., 2022. Corporate ethical behaviours and firm equity value and ownership: evidence from the gpfg's ethical exclusions. Available at SSRN 3388868 .

Berk, J.B., Van Binsbergen, J.H., 2025. The impact of impact investing. Journal of Financial Economics 164, 103972.

Berle, E., He, W.A., Ødegaard, B.A., 2025. The stock market and corporate consequences of ethical exclusions by the world's largest fund. Journal of International Financial Markets, Institutions and Money 102, 102174.

Białkowski, J., Gottschalk, K., Wisniewski, T.P., 2008. Stock market volatility around national elections. Journal of banking & finance 32, 1941–1953.

Blitz, D., Fabozzi, F.J., 2017. Sin stocks revisited: Resolving the sin stock anomaly .

Boehmer, E., Masumeci, J., Poulsen, A.B., 1991. Event-study methodology under conditions of event-induced variance. Journal of financial economics 30, 253–272.

Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. Journal of econometrics 31, 307–327.

Bolton, P., Kacperczyk, M., 2021. Do investors care about carbon risk? Journal of financial economics 142, 517–549.

Chava, S., 2014. Environmental externalities and cost of capital. Management science 60, 2223–2247.

Dimson, E., Karakaş, O., Li, X., 2015. Active ownership. The Review of Financial Studies 28, 3225–3268.

Edmans, A., Levit, D., Schneemeier, J., 2022. Socially responsible divestment. Centre for Economic Policy Research.

Engle, R.F., 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. Econometrica: Journal of the econometric society , 987–1007.

Glosten, L.R., Jagannathan, R., Runkle, D.E., 1993. On the relation between the expected value and the volatility of the nominal excess return on stocks. The journal of finance 48, 1779–1801.

Hartzmark, S.M., Shue, K., 2022. Counterproductive sustainable investing: The impact elasticity of brown and green firms. Available at SSRN 4359282 .

Hartzmark, S.M., Sussman, A.B., 2019. Do investors value sustainability? a natural experiment examining ranking and fund flows. The Journal of Finance 74, 2789–2837.

Heinkel, R., Kraus, A., Zechner, J., 2001. The effect of green investment on corporate behavior. Journal of financial and quantitative analysis 36, 431–449.

Hoepner, A.G., Schopohl, L., 2018. On the price of morals in markets: An empirical study of the swedish ap-funds and the norwegian government pension fund. Journal of Business Ethics 151, 665–692.

Hong, H., Kacperczyk, M., 2009. The price of sin: The effects of social norms on markets. Journal of financial economics 93, 15–36.

Luo, H.A., Balvers, R.J., 2017. Social screens and systematic investor boycott risk. Journal of financial and quantitative analysis 52, 365–399.

MacKinlay, A.C., 1997. Event studies in economics and finance. Journal of economic literature 35, 13–39.

Merton, R.C., et al., 1987. A simple model of capital market equilibrium with incomplete information .

Nelson, D.B., 1991. Conditional heteroskedasticity in asset returns: A new approach. Econometrica: Journal of the econometric society , 347–370.

Nguyen, P.A., Kecskés, A., Mansi, S., 2020. Does corporate social responsibility create shareholder value? the importance of long-term investors. Journal of Banking & Finance 112, 105217.

Pástor, L., Stambaugh, R.F., Taylor, L.A., 2021. Sustainable investing in equilibrium. Journal of financial economics 142, 550–571.

Pedersen, L.H., Fitzgibbons, S., Pomorski, L., 2021. Responsible investing: The esg-efficient frontier. Journal of financial economics 142, 572–597.

Rudolf, N., Yuan, H., 2025. Investor alignment in divestment decisions and firm behavior: Evidence from publicly disclosed exclusion lists. European Accounting Review , 1–27.

Savickas, R., 2003. Event-induced volatility and tests for abnormal performance. Journal of Financial Research 26, 165–178.

Teoh, S.H., Welch, I., Wazzan, C.P., 1999. The effect of socially activist investment policies on the financial markets: Evidence from the south african boycott. The Journal of Business 72, 35–89.

Xu, Z., Liu, D., Li, Y., Guo, F., 2025. Esg and stock price volatility risk: Evidence from chinese a-share market. The North American Journal of Economics and Finance 75, 102277.

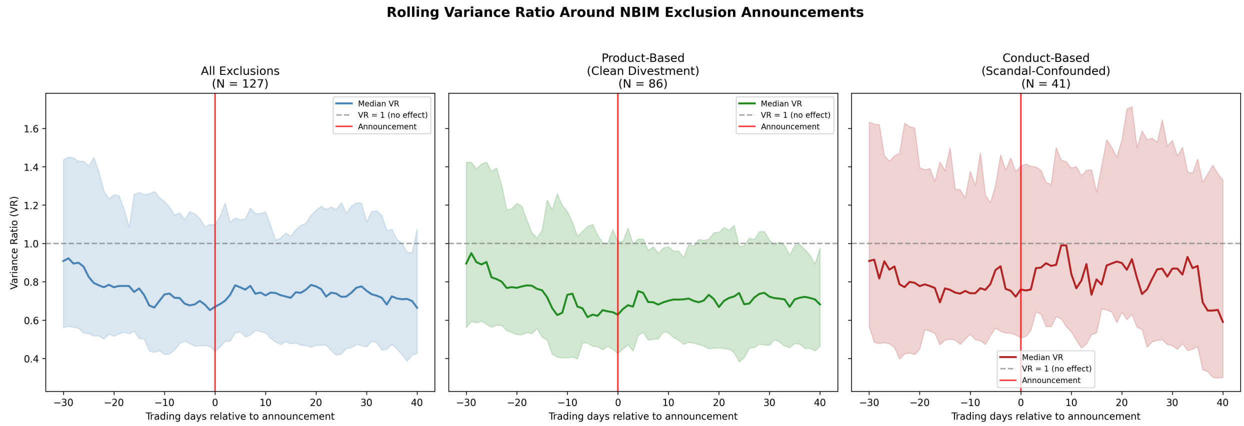


**Figure 1:** Rolling Variance Ratio Around NBIM Exclusion Announcements

*Notes*: Each panel shows the cross-sectional average of firm-level rolling variance ratios (20-day rolling window, normalized by estimation-period variance) around the exclusion announcement date (day 0). Shaded area represents the $[-10, +20]$ event window. Left panel: all exclusions. Center: product-based exclusions (tobacco, weapons, coal). Right: conduct-based exclusions (human rights, corruption, environment). The horizontal dashed line at VR $= 1$ indicates no abnormal variance.

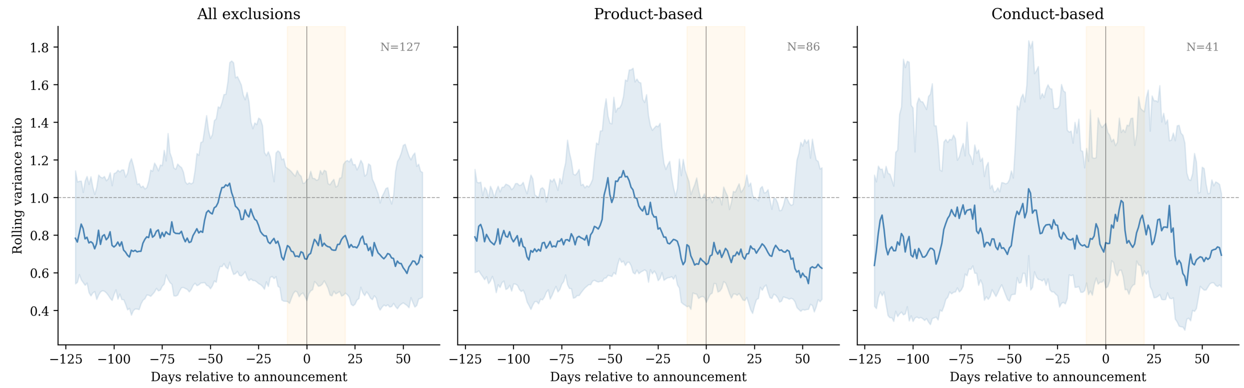


**Figure 2:** Extended Rolling Variance Ratio: $[-180, +60]$ Days

*Notes*: Each panel plots the cross-sectional median of firm-level 20-day rolling variance ratios from 180 trading days before to 60 trading days after the exclusion announcement. Shaded band: interquartile range. Orange shading marks the main $[-10, +20]$ event window. Day 0 is the announcement date. The baseline variance is estimated from a 252-day window ending at day $-181$.

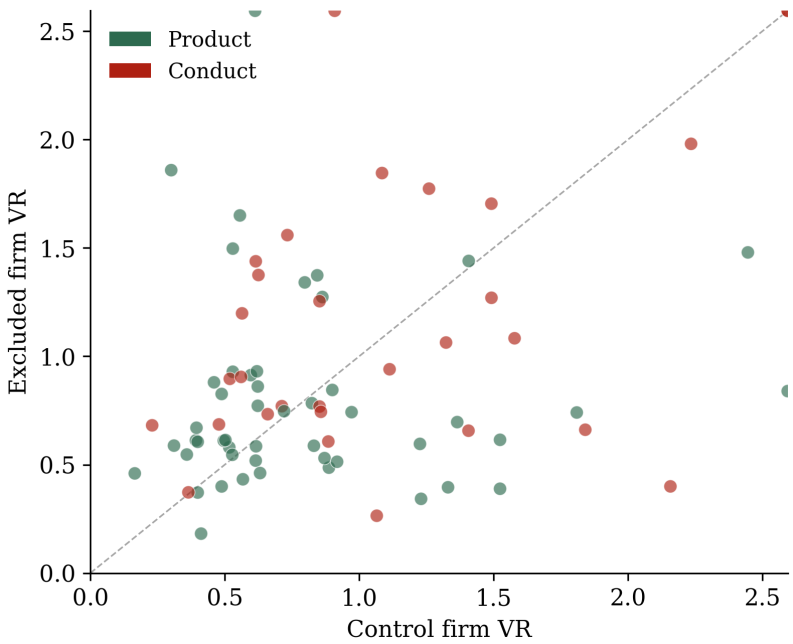


**Figure 3:** Variance Ratios: Excluded Firms vs. Matched Controls

*Notes*: Each point represents a matched pair. The $x$-axis shows the control firm's variance ratio and the $y$-axis the excluded firm's variance ratio, both computed around the exclusion announcement date. Points on the 45-degree dashed line indicate identical volatility behavior. Green: product exclusions; red: conduct exclusions.

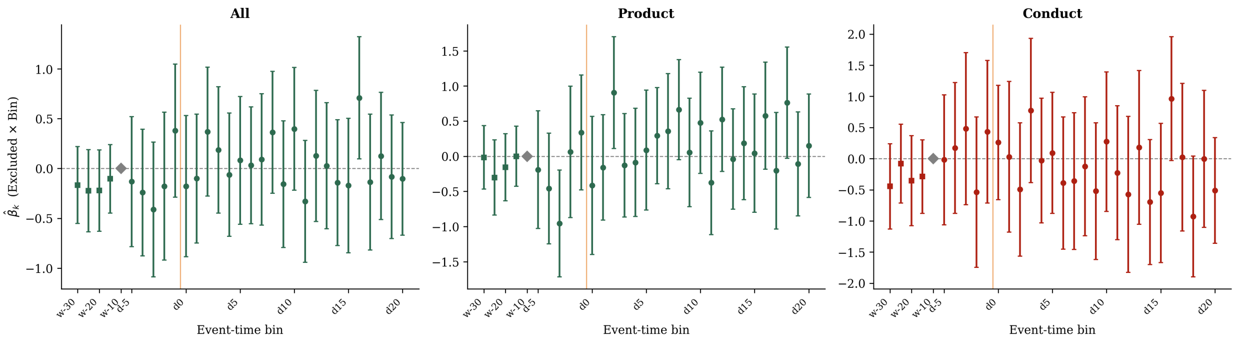


**Figure 4:** Panel DiD Event-Time Coefficients

*Notes*: Each panel plots $\hat{\beta}_k$ from a regression of $\log(\mathrm{RV}_{i,t})$ on Excluded$_i \times D_k$ (event-time bin dummies) with firm fixed effects. Pre-period bins are weekly ($[-30, -26]$, $[-25, -21]$, etc.); near-event and post bins are daily. The reference bin is $[-10, -6]$ (coefficient normalised to zero). Shaded bands are 95% confidence intervals based on firm-clustered standard errors. Squares denote weekly bins; circles denote daily bins.

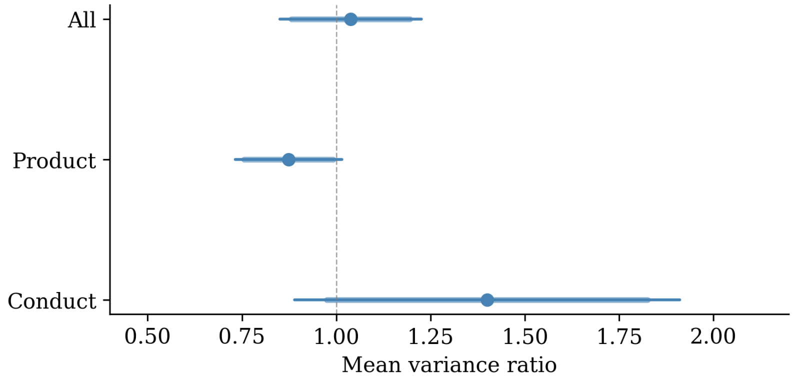


**Figure 5:** Confidence Intervals for Mean Variance Ratio

*Notes*: Forest plot showing 90% (thick) and 95% (thin) confidence intervals for the cross-sectional mean VR. Circle: point estimate. Dashed line at VR = $1$.

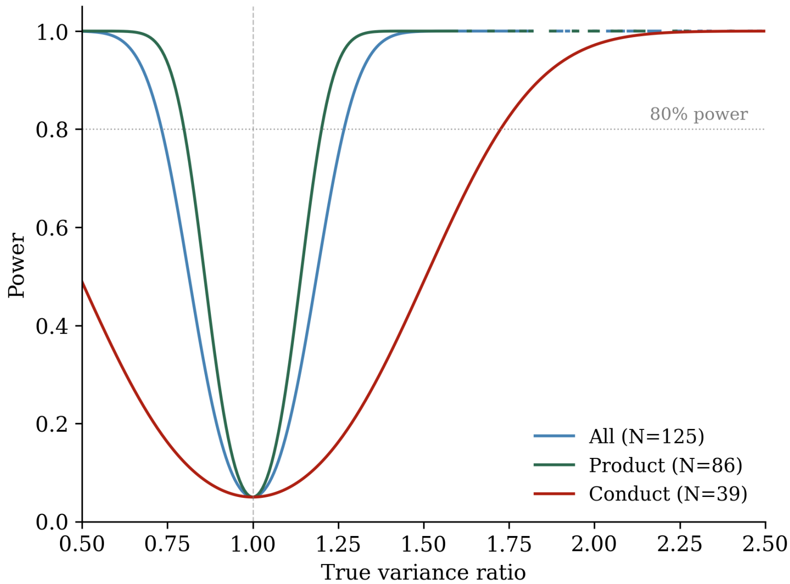


**Figure 6:** Power Curves by Category

*Notes*: Power (probability of rejecting $H_0$: VR = $1$ at $\alpha = 0.05$) as a function of the true variance ratio. Horizontal dashed line: 80% power threshold. Vertical dashed line at VR = $1$.