sustainability

MDPI

Article

Defence-Sector Intensity and Foreign Direct Investment in the European Union Under Global Uncertainty: A Multi-Method Assessment

Tomas Ruzgas [1,*] and Smiltė Ruzgaitė [2]

[1] Department of Applied Mathematics, Kaunas University of Technology, LT-51368 Kaunas, Lithuania
[2] Faculty of Economics and Management, Vytautas Magnus University, LT-44244 Kaunas, Lithuania
* Correspondence: tomas.ruzgas@ktu.lt

Abstract

This study examines the association between defence-sector intensity and foreign direct investment (FDI) in 23 European Union countries from 2005 to 2023. Six defence indicators and three global uncertainty measures were analysed using fixed-effects panel regressions, regularised regression, gradient boosting, country-profile comparisons, and panel temporal-predictability tests. In the principal country-and-year fixed-effects models, military expenditure as a percentage of GDP and as a percentage of government expenditure had positive contemporaneous coefficients. Both coefficients remained statistically supported by Driscoll–Kraay and wild-cluster-bootstrap inference after adjustment for the two primary hypotheses, although their Holm-adjusted country-clustered *p*-values were 0.053. The findings were sensitive to country coverage: excluding Hungary, Ireland, and the Netherlands reduced the standardised estimates, and neither coefficient remained statistically significant under wild-cluster-bootstrap inference. Neither primary indicator was supported in the available-case EU27 extension; matched-period comparisons indicated that this divergence reflected country composition together with the shorter common observation window. Defence indicators did not provide stable incremental predictive information for countries excluded from model fitting, and no temporal-predictability result remained significant after dependence-aware block-bootstrap inference and false-discovery-rate adjustment. Neither global geopolitical risk nor the country profiles provided evidence of coefficient differences. Overall, the evidence is indicator-, sample-, and method-dependent and does not establish a causal or generally applicable effect of defence expenditure on FDI. From a sustainable-development perspective, security expenditure should therefore be assessed alongside fiscal capacity, investment conditions, and country-specific economic structures rather than being assumed either to attract or to displace foreign investment uniformly.

Keywords: foreign direct investment; military expenditure; defence-sector intensity; European Union; sustainable economic development; fixed-effects panel regression; cross-sectional dependence; predictive validation; temporal predictability

Academic Editor: Izabela Jonek-Kowalska

Received: 18 September 2026
Revised: 26 October 2026
Accepted: 27 October 2026
Published: 28 October 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

1.1. Defence-Sector Intensity, Economic Resilience, and Foreign Direct Investment

Changes in the European security environment have increased the fiscal and economic importance of defence expenditure. In 2023, the 27 European Union (EU) Member States spent EUR 279 billion on defence, 10% more in real terms than in 2022. Defence investment reached EUR 72 billion and represented 26% of total defence expenditure [1]. Security-related spending may strengthen deterrence, industrial capacity, technological development, and economic resilience, but it may also compete with infrastructure, education, and other long-term priorities. Its relevance to sustainable development therefore lies not in assuming that defence expenditure is inherently sustainable but in examining how security demands interact with economic capacity, public-resource allocation, and capital formation. These effects may differ across countries because their fiscal capacity, institutional quality, industrial structure, and security exposure are not uniform [2].

Foreign direct investment is relevant to this relationship because it connects host economies with long-term capital, technology, managerial knowledge, and international production networks. Under the OECD definition, FDI is a cross-border investment that gives an investor a lasting interest and significant influence over an enterprise in another economy, conventionally indicated by ownership of at least 10% of its voting power [3]. Stable international investment can support productive capacity and longer-term economic development. Nevertheless, the outcome used in this study—annual net FDI inflows as a percentage of GDP—does not distinguish green, socially responsible, or otherwise sustainability-oriented investment. It records realised net financial flows that may include equity investment, reinvested earnings, intracompany financing, corporate restructuring, pass-through transactions, and disinvestment.

The expected relationship between defence-sector intensity and FDI is theoretically ambiguous. Greater defence capacity may coincide with deterrence, perceived stability, technological demand, or participation in international defence-production networks, thereby supporting economic resilience and investment conditions. Higher military expenditure may instead reflect geopolitical exposure, create fiscal pressure, or displace expenditure that contributes to productive and social development. The estimated relationship may consequently depend on the defence indicator, control set, functional form, country coverage, and temporal specification. Global uncertainty may further influence investment decisions because geopolitical and policy shocks can encourage firms to postpone investments involving substantial sunk costs [4]. Pooled correlations, out-of-country predictive performance, conditional panel associations, and lagged predictability address different statistical questions and may therefore yield different conclusions. In this study, these forms of evidence are interpreted separately, and none is treated as identifying a causal effect.

1.2. Related Research and Research Gap

Research on the determinants of FDI identifies several recurring host-country characteristics. In a systematic review of 112 empirical studies, Islam and Beloucif [5] found the strongest support for market size, followed by trade openness, infrastructure quality, labour costs, macroeconomic stability, human capital, and growth prospects. Political and institutional conditions are also important because many foreign investments involve substantial sunk costs and long planning horizons. By expanding productive capacity, facilitating technology transfer, and connecting host economies to international production networks, FDI may contribute to longer-term development and economic resilience. However, the volume of FDI alone does not establish whether the underlying investments are environmentally or socially sustainable, and this study does not attempt to measure those

characteristics. Related evidence from environmental-finance research likewise indicates that relationships among investment, geopolitical-economic uncertainty, and sustainability outcomes may vary across model specifications and country groups [6].

Evidence on the relationship between defence expenditure and FDI is more limited and does not indicate a common direction. Aziz and Khalid [7] found that military expenditure reduced FDI inflows in developing countries without armed conflict, while conflict exposure altered this relationship and the estimated response emerged over time. Douch [8], by contrast, reported a positive relationship between defence expenditure and FDI in Middle Eastern and North African countries using a dynamic panel model. Bhandari et al. [9] further showed that conflict dynamics and institutional conditions can modify the military-expenditure–FDI relationship. These findings cannot be transferred directly to the EU, where institutional arrangements, security commitments, fiscal capacity, industrial structures, and exposure to armed conflict differ from those of the countries examined in earlier studies.

The literature also leaves unresolved problems of measurement and interpretation. Defence-sector intensity is often represented by aggregate military expenditure, although expenditure relative to GDP, budgetary allocation, per-capita expenditure, defence investment, militarisation, and arms imports describe different economic features. A relationship found for one measure need not extend to the others. The methods used to examine these relationships also answer different questions. Fixed-effects panel models estimate conditional within-country associations, regularised regression screens correlated predictors, and nonlinear algorithms examine predictive patterns that may not be captured by a linear specification. Their results are not interchangeable: predictor retention does not establish statistical significance, and in-sample model fit does not demonstrate that predictions will transfer to countries excluded from estimation. Lagged panel tests address a further question by examining temporal predictability rather than the magnitude or causal direction of an effect.

Few studies have compared several measures of defence-sector intensity within a common EU panel while distinguishing contemporaneous association, predictive relevance, out-of-country performance, and temporal predictability. Evidence is particularly limited on whether defence indicators remain informative after economic and institutional controls are introduced, whether predictive models generalise to countries excluded from model fitting, and whether estimated associations are sensitive to influential countries, extreme FDI observations, alternative scaling, and wider country coverage. It is also unclear whether global geopolitical risk or data-derived country profiles systematically modify the defence–FDI relationship. These questions are relevant to sustainable economic development because changes in security expenditure may coincide with changes in fiscal priorities, investment conditions, and the capacity to maintain international capital formation. The present study addresses these gaps using annual observations for 23 EU countries from 2005 to 2023, supported by sensitivity analyses based on alternative samples and variable definitions.

1.3. Research Aim, Questions, and Contributions

This study examines whether the relationship between defence-sector intensity and FDI in the European Union is consistent across indicators, model specifications, and statistical questions. The analysis covers 23 EU countries from 2005 to 2023 and represents defence-sector intensity using six indicators: military expenditure relative to GDP, military expenditure relative to government expenditure, military expenditure per capita, defence investment, the Global Militarisation Index, and arms imports. Principal component analysis, clustering, regularised regression, gradient-boosted trees, fixed-effects panel

regression, and panel temporal-predictability tests are used for distinct analytical purposes rather than as interchangeable estimators.

The study addresses four research questions:

- RQ1. Which defence-sector intensity indicators exhibit conditional contemporaneous associations with FDI, and how sensitive are the primary estimates to the control structure, inference method, influential observations, variable scaling, and country coverage?
- RQ2. Which defence-sector indicators provide stable linear or nonlinear predictive information for FDI, and do the resulting models perform better than a training-mean benchmark when applied to countries excluded from model fitting?
- RQ3. Do the conditional associations between defence-sector intensity and FDI vary systematically with geopolitical risk or across data-derived country profiles?
- RQ4. Do defence-sector indicators and FDI exhibit panel-level temporal predictability, and is the identified predictive information unidirectional or bidirectional?

The study makes three principal contributions. First, it compares six non-interchangeable measures of defence-sector intensity in a common EU panel instead of relying only on military expenditure relative to GDP. Two expenditure-share measures are designated as primary indicators, while the remaining measures provide secondary evidence concerning expenditure intensity, investment, militarisation, and arms transfers. This distinction reduces the risk of treating results selected from several indicators as equivalent confirmatory findings.

Second, the analysis separates explanatory and predictive questions. The panel regressions use a prespecified B0–B2 control hierarchy and combine country-clustered, Driscoll–Kraay, and wild-cluster-bootstrap inference. Sensitivity analyses examine extreme FDI observations, influential countries, alternative scaling, and an available-case EU27 extension. Predictive performance is evaluated through nested country-grouped cross-validation, so observations from a validation country are excluded from model fitting and preprocessing. Country-profile and geopolitical-risk interactions are assessed using formal joint tests rather than comparisons of separately estimated subgroup coefficients.

Third, the study relates the empirical evidence to sustainable economic development without treating either defence expenditure or aggregate FDI as a direct measure of sustainability. The results indicate whether relationships between security-related resource intensity and international capital inflows remain stable across measures and analytical procedures. This is relevant when evaluating security requirements alongside fiscal priorities, productive investment, and longer-term economic resilience. Because the study is observational, the regression, predictive, and temporal results are interpreted as conditional associations or evidence of predictability rather than as structural causal effects.

2. Materials and Methods

2.1. Analytical Strategy

The analysis was organised around a common panel-regression framework. The primary analysis examined two measures of defence-sector intensity: military expenditure as a percentage of GDP and military expenditure as a percentage of government expenditure. Both measures were estimated using the same country and year fixed effects and the same sequence of control sets. This common structure allowed the estimated associations to be compared without changes in the underlying fixed-effects specification.

Three nested specifications were used. Model B0 included the defence indicator together with country and year fixed effects. Model B1 added macroeconomic controls: GDP per capita, trade openness, inflation, gross capital formation, and unemployment. Model B2 additionally controlled for secondary-school enrolment, research and development

expenditure, and regulatory quality. Model B2 was treated as the principal specification because it provided the most detailed adjustment while retaining almost the complete panel. The B0–B2 hierarchy and the designation of the two primary indicators were fixed before the revised estimation undertaken in response to peer review, but after the original analysis had been inspected. They organise the reanalysis and should not be interpreted as a prospectively registered confirmatory design.

Military expenditure per capita, defence investment, the Global Militarisation Index, and arms imports were examined as secondary defence indicators. Defence investment and military expenditure per capita were log-transformed, while arms imports were transformed using the inverse hyperbolic sine. Originally missing arms-import observations remained missing in the principal analysis. Geopolitical-risk moderation, country-profile differences, regularised regression, machine-learning models, and temporal-predictability tests were treated as exploratory analyses addressing questions distinct from the contemporaneous panel associations.

Robustness was assessed using country-clustered covariance estimates, Driscoll–Kraay standard errors, and wild-cluster bootstrap inference. Additional checks included alternative Driscoll–Kraay bandwidths, winsorisation of FDI, leave-one-country-out estimation, joint exclusion of Hungary, Ireland, and the Netherlands, an alternative per-capita formulation intended to reduce the common-denominator problem, period-based restrictions, and an available-case analysis covering all 27 EU Member States. Principal component analysis and clustering were retained for descriptive country profiling rather than used as the principal means of controlling for observed economic and institutional conditions. Predictive models were evaluated using nested cross-validation grouped by country, with all preprocessing and tuning repeated within the corresponding training samples.

The procedures estimate conditional associations and predictive patterns rather than structural causal effects. A result was considered comparatively robust only when its sign remained stable and the principal inference procedures and sensitivity analyses led to similar conclusions. Results that depended on the country sample, outcome scaling, functional form, or estimation method were classified as specification-sensitive and interpreted accordingly.

2.2. Data Sources and Study Sample

The source data initially covered all 27 EU Member States from 1995 to 2024. FDI and most control variables were obtained from World Bank Open Data and Eurostat [10,11]. Military-expenditure indicators came primarily from the SIPRI Military Expenditure Database, while defence-investment data were obtained from the European Defence Agency's Defence Data portal [12,13]. The sources of the remaining variables are reported in Section 2.3.

FDI, defence, and control variables were available annually. Monthly observations of the Global Economic Policy Uncertainty and Geopolitical Risk indices and quarterly observations of the World Uncertainty Index were aggregated into annual arithmetic means. After country and year identifiers had been harmonised, the source datasets were merged into an initial panel containing 810 country–year observations and 47 analytical variables, excluding the identifiers.

Because most missing observations occurred before 2005 or in 2024, the study period was restricted to 2005–2023. The principal analytical sample comprised 23 countries: Austria, Belgium, Bulgaria, Croatia, Czechia, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, the Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, and Sweden. Cyprus, Denmark, Luxembourg, and Malta were not included in this sample because their coverage was insufficient for parts of the broader

variable set used in the secondary, principal component, clustering, and predictive analyses. Their exclusion did not imply that the two primary military-expenditure indicators were entirely unavailable.

The principal sample contained 437 country–year observations, corresponding to 23 countries observed over 19 years. The country–year structure was balanced, although 14 defence-investment observations and one secondary-school-enrolment observation remained unavailable after data preparation. Analyses involving these variables used the observations available for the corresponding specification. To assess whether restricting the sample to 23 countries affected the principal panel-regression results, the B2 models for the two primary military-expenditure indicators were additionally estimated for all 27 EU Member States using available observations. This sensitivity sample contained 507 country–year observations: 18 for Cyprus, 18 for Ireland, 15 for Luxembourg, and 19 for each of the other 24 countries. No values were extrapolated solely to complete the EU27 sensitivity sample.

2.3. Variables and Measurement

The analysis distinguished FDI as the outcome, two primary defence indicators, four secondary defence indicators, three global uncertainty measures, and 37 initially considered controls. Table 1 summarises their measurement and principal sources.

Table 1. Variable groups and operationalisation.

Variable Group	Indicators	Measurement and Principal Source
Outcome	Foreign direct investment	Net FDI inflows as a percentage of GDP; World Bank Open Data
Primary defence indicators	Military expenditure as a percentage of GDP; military expenditure as a percentage of government expenditure	Percentage of GDP and percentage of government expenditure; SIPRI Military Expenditure Database
Secondary defence intensity	Military expenditure per capita; defence investment; Global Militarisation Index; arms imports	EUR per capita, EUR million, composite index score, and SIPRI trend-indicator values; SIPRI, European Defence Agency, BICC, and World Bank Open Data
Uncertainty	Global Economic Policy Uncertainty (GEPU) Index; Geopolitical Risk Index; World Uncertainty Index	Monthly or quarterly global indices aggregated into annual arithmetic means; respective index database
Economic controls	Initially 12 indicators	Economic scale and performance, labour costs, capital formation, exchange-rate conditions, and economic freedom; World Bank Open Data, Eurostat, and institutional databases
Infrastructure controls	Initially seven indicators	Energy, digital, transport, and research infrastructure; World Bank Open Data and Eurostat
Political-institutional controls	Thirteen indicators	Governance, political stability, institutional constraints, taxation, and regulatory and economic freedom; World Bank Open Data and institutional databases
Human-capital controls	Initially five indicators	Population, education, labour-market conditions, and tourist arrivals; World Bank Open Data and Eurostat

Net FDI inflows were divided by GDP to improve comparability across economies of different sizes. The resulting variable measures realised net financial flows rather than investment attractiveness, as discussed in Section 1.1.

Defence-sector intensity was measured from several perspectives. Military expenditure as a percentage of GDP and as a percentage of government expenditure were designated as the primary indicators because they measure, respectively, the economic burden of defence and its share of public expenditure. Military expenditure per capita, defence investment, the Global Militarisation Index (GMI), and arms imports were examined as secondary indicators. Defence investment covered equipment procurement and research and development expenditure classified by the European Defence Agency as investment. The GMI, produced by the Bonn International Centre for Conflict Studies (BICC), combines military expenditure, personnel, and heavy weapons relative to economic, social, and demographic benchmarks [14]; it does not measure warfare or conflict intensity. Arms imports were expressed as SIPRI trend-indicator values distributed through World Bank Open Data. These values measure the volume of international transfers of major conventional weapons rather than their physical number or financial value.

The Global Economic Policy Uncertainty (GEPU) Index measures the relative frequency of newspaper articles referring to the economy, policy, and uncertainty [15]. The Geopolitical Risk Index uses newspaper coverage of wars, terrorist acts, international tensions, and related threats [4], while the World Uncertainty Index is derived from uncertainty-related language in Economist Intelligence Unit country reports [16]. Because all three series were global, every country received the same value in a given year. They therefore represent common annual uncertainty conditions rather than cross-country differences in exposure.

The initial controls comprised 12 economic, seven infrastructural, 13 political-institutional, and five human-capital indicators. Two economic variables were removed during data-availability screening: average annual earnings, which were unavailable for the Netherlands throughout 2005–2023, and government bond yields, which were unavailable for Estonia throughout 2005–2019. Retaining them would have required unsupported cross-country imputation. The principal B2 panel specification used eight observed controls selected to represent the main economic and institutional conditions associated with FDI: GDP per capita, trade openness, inflation, gross capital formation, unemployment, secondary-school enrolment, research and development expenditure, and regulatory quality. The use of observed controls in the principal specification avoided relying on the PCA indices for regression adjustment.

2.4. Data Preparation and Preliminary Analysis

Country names, ISO codes, and year identifiers were harmonised before merging. Missingness was assessed by year, country, and variable. Interpolation was confined to internal gaps within national time series; values were neither interpolated across countries nor extrapolated beyond observed series boundaries.

Thirteen missing military-expenditure-per-capita values were calculated from data for the same country and year using

$$MEPC_{i,t} = \frac{MES_{i,t}}{100} \times \frac{GDP_{i,t} \times 1000}{POP_{i,t}},$$

where $MEPC_{i,t}$ denotes military expenditure per capita in EUR, $MES_{i,t}$ is military expenditure expressed as a percentage of GDP, $GDP_{i,t}$ is GDP in EUR million, and $POP_{i,t}$ is population in thousands.

Cubic-spline interpolation completed five secondary-school-enrolment values and three tourist-arrival values. The 40 originally missing arms-import observations remained missing rather than being interpreted as verified zero transfers. Analyses involving arms imports therefore used the observations for which SIPRI trend-indicator values were originally available. No arms-import values were generated through interpolation or zero substitution.

Descriptive analysis comprised the mean, median, standard deviation, interquartile range, skewness, and kurtosis. Annual medians and interquartile ranges were plotted for FDI and the two primary military-expenditure indicators to show their evolution over 2005–2023. The years 2014 and 2022 were marked as potentially relevant changes in the European security environment, while 2020–2021 was identified as the main COVID-19 period. These markers were used for descriptive orientation and did not by themselves define causal events. Potentially influential country–year observations were additionally assessed using studentised residuals, leverage, and Cook's distance in the principal B2 models. Observations were not removed solely because their values were extreme, as unusually large positive or negative FDI flows may reflect genuine economic events.

The robustness analysis examined whether the principal B2 results depended on extreme observations, country coverage, the study period, or variable scaling. First, FDI was winsorised at the pooled 1st and 99th percentiles. Second, each country was omitted in turn. Third, Hungary, Ireland, and the Netherlands were excluded jointly because these countries accounted for several of the largest positive or negative FDI observations and exerted appreciable influence in the diagnostic analysis. Fourth, the models were re-estimated for 2005–2021 to exclude the period following the 2022 escalation of the war in Ukraine. Fifth, observations from 2020 and 2021 were removed to assess sensitivity to the main COVID-19 period. Sixth, the common-GDP-denominator concern was examined by expressing both FDI and military expenditure on a per-capita basis and applying transformations appropriate for negative or strongly right-skewed values. Finally, the two primary B2 models were estimated for all 27 EU Member States using the available observations. To distinguish the influence of country composition from that of data availability, the EU27 dataset was also restricted first to the original 23 countries and then to the 14 years for which complete observations were available for all 27 countries. Corresponding models for the original 23-country sample were estimated over the same 14-year period. Each sensitivity model retained the B2 controls and country and year fixed effects. Country-clustered, Driscoll–Kraay, and wild-cluster-bootstrap inference was used where the corresponding sample structure permitted it.

The LLC and IPS tests were retained as first-generation stationarity diagnostics [17,18]. Because EU country series may respond to common economic, financial and geopolitical shocks, the principal stationarity assessment used Pesaran's cross-sectionally augmented IPS (CIPS) test, which augments country-specific Dickey–Fuller regressions with cross-sectional averages [19]. The regressions included an intercept, and country-specific lag orders from zero to two were selected using the Bayesian information criterion. Statistical decisions were based on the published finite-sample critical values nearest to the dimensions of the corresponding panel. The CIPS tests were conducted for the complete 23-country sample, the reduced sample excluding Hungary, Ireland, and the Netherlands, and, for FDI, the 18-country and 13-country balanced samples used in the temporal tests involving defence investment and arms imports, respectively. Country-specific augmented Dickey–Fuller and KPSS statistics were additionally examined as descriptive diagnostics of within-country heterogeneity. Given the short national series, these country-level tests were not used as independent binary classifications of every country's data-generating process. The complete panel and country-level stationarity diagnostics are reported in Table S1A–C, while the first-difference sensitivity model is reported in Table S1D.

In the complete 23-country sample, the CIPS test rejected the unit-root null at the 5% level for FDI, both primary military-expenditure shares, log defence investment, GMI, and the inverse-hyperbolic-sine-transformed arms-import series. For log military expenditure per capita, rejection was obtained only at the 10% level. After Hungary, Ireland, and the Netherlands were excluded, the unit-root null was rejected at the 5% level for FDI, both

primary expenditure shares, log defence investment, GMI, and transformed arms imports, but it was not rejected for log military expenditure per capita even at the 10% level. For FDI in the smaller balanced samples used in the temporal analyses, the unit-root null was rejected at the 1% level in the 18-country defence-investment sample (CIPS = −2.731), but only at the 10% level in the 13-country arms-import sample (CIPS = −2.116). The latter result was therefore treated as weaker sample-specific evidence, and the temporal findings involving arms imports were interpreted as exploratory. The country-specific ADF and KPSS diagnostics likewise indicated heterogeneous time-series properties across countries and variables. Accordingly, the panel-level rejections were interpreted as evidence concerning the corresponding panels rather than as proof that every national series was stationary. Log military expenditure per capita was first-differenced for the temporal-predictability analysis:

$$\Delta x_{i,t} = x_{i,t} - x_{i,t-1}.$$

The remaining variables entered the temporal-predictability tests in levels. As an additional sensitivity check, the association between changes in log military expenditure per capita and changes in FDI was estimated using a first-difference version of the fully adjusted B2 specification. The contemporaneous fixed-effects regressions otherwise used the level or transformed forms defined for their respective specifications and were interpreted as conditional same-period associations rather than as long-run equilibrium relationships.

Pearson correlations between FDI and the defence, uncertainty, and control variables were calculated using the observations available for each pair. Within-group correlations among controls were also examined for strong pairwise dependencies. As these pooled correlations did not control for country effects or serial dependence, they were treated as exploratory. Unadjusted p-values below 0.05 were reported descriptively but were not used as the sole criterion for retaining predictors.

2.5. Dimensionality Reduction and Country Profiling

PCA was applied separately to the economic, infrastructural, political-institutional, and human-capital variable groups to construct descriptive indices for country profiling [20]. PCA was performed on pooled country–year observations and therefore did not model the dependence between repeated observations from the same country. For this reason, and because the adequacy of a single-component representation differed across the four groups, the resulting indices were not used as controls in the principal B2 panel regressions.

$$z_{i,t,j} = \frac{x_{i,t,j} - \overline{x_j}}{s_j},$$

where $\overline{x_j}$ and s_j are the mean and standard deviation of variable j. The analysis was consequently based on correlations rather than on the original measurement scales.

Sampling adequacy was evaluated using the overall Kaiser–Meyer–Olkin (KMO) measure and individual measures of sampling adequacy [21]. Hourly labour costs and internet use were removed from the economic and infrastructural groups, respectively, because of inadequate individual values. Tourist arrivals were removed from the human-capital group because of their strong redundancy with population. No political-institutional variable was removed.

The final groups contained nine economic, six infrastructural, 13 political-institutional, and four human-capital indicators, with overall KMO values of 0.559, 0.786, 0.887, and 0.480, respectively. The human-capital group was retained despite falling below the conventional 0.50 threshold because its variables formed a theoretically coherent set; the resulting component was interpreted cautiously.

The number of components supported by the data was examined using parallel analysis with 1000 independently permuted datasets. For each component, the observed eigenvalue was compared with the 95th percentile of the corresponding eigenvalues obtained from the permuted data. The parallel analysis supported three components for the economic group, one for infrastructure, two for the political-institutional group, and one for human capital. The corresponding PCA diagnostics are reported in Section 3.2, while the complete PCA summary and PC1 loadings are provided in Table S2A and S2B, respectively.

For country profiling, PC1 explained 25.1% of the standardised variance in the economic group, 65.4% in infrastructure, 51.1% in the political-institutional group, and 36.0% in human capital. The relatively low proportions for the economic and human-capital groups, together with the parallel-analysis results, limit the interpretation of these scores as complete representations of their respective domains. As component signs are arbitrary, the scores were interpreted from their joint loading patterns rather than from their absolute orientation.

For each country, the four indices were averaged over 2005–2023. Each of the four resulting country-level variables was then standardised across the 23 countries. The final 23×4 matrix therefore contained standardised country means rather than the original pooled PC1 scores. FDI, defence indicators, and uncertainty indices were not used to construct the profiles.

Six clustering methods were compared: K-means, spectral clustering, Ward hierarchical clustering, complete-linkage agglomerative clustering, Gaussian mixture modelling, and affinity propagation. Solutions containing two to six clusters were assessed for methods requiring the number of clusters to be specified, while affinity propagation selected this number algorithmically. Internal validation used the silhouette coefficient, Davies–Bouldin index, and Calinski–Harabasz statistic [22–24]. Table S3 reports the complete validation results.

The three-cluster K-means solution had the highest silhouette coefficient among the K-means candidates and was fitted using 50 random initialisations and a fixed random seed of 12,345,678. Ward hierarchical clustering reproduced the same partition. The standardised country-level matrix and the resulting cluster centres are reported to make the profile construction reproducible, while Figure S1 presents a heatmap of the four standardised profile dimensions and the corresponding cluster membership of each country. The profiles were treated as descriptive groupings. Differences in defence–FDI coefficients were evaluated through pooled interaction models rather than inferred from statistical significance within separately estimated profile-specific regressions.

2.6. Regularised Regression and Gradient Boosting

Regularised regression and gradient-boosted regression trees were used as supplementary predictive analyses. Ridge, Lasso, and Elastic Net assessed linear predictive patterns under coefficient penalisation, while gradient boosting allowed nonlinearities and higher-order dependencies [25]. These models were evaluated according to their ability to predict FDI for countries excluded from model fitting and were not treated as causal estimators or as substitutes for the panel regressions.

Ridge, Lasso, and Elastic Net minimised

$$\left(\hat{\beta}_0, \hat{\boldsymbol{\beta}}\right) = \arg\min_{\beta_0, \boldsymbol{\beta}} \left[\frac{1}{2n} \sum_{i=1}^{n} \left(y_i - \beta_0 - \mathrm{x}_i^{\mathrm{T}} \boldsymbol{\beta}\right)^2 + \lambda \left(\alpha \sum_{j=1}^{p} |\beta_j| + \frac{1-\alpha}{2} \sum_{j=1}^{p} \beta_j^2 \right) \right]$$

where n is the number of observations, p is the number of predictors, y_i is the FDI value, and x_i is the predictor vector for observation i. The coefficient vector is $\boldsymbol{\beta}$, $\lambda \geq 0$

controls penalty strength, and $\alpha \in [0, 1]$ determines the balance between the L_1 and L_2 penalties. The intercept β_0 was not penalised. Ridge corresponded to $\alpha = 0$ and Lasso to $\alpha = 1$, while the Elastic Net tuning procedure considered α values of 0.25, 0.50, and 0.75 [26–28]. The overall penalty parameter λ was selected within the inner cross-validation loop.

The principal predictive dataset contained the six defence indicators, three global uncertainty measures, and one PCA score for each of the four control-variable groups. Observations without originally recorded arms-import or defence-investment values were excluded, leaving 385 country–year observations. FDI remained expressed as a percentage of GDP. To assess the incremental predictive contribution of the defence indicators, a corresponding reduced predictor set was defined by removing all six defence variables while retaining the uncertainty measures and the four PCA-based control scores. The predictive analysis was kept separate from the principal B2 panel models, which used the eight observed controls defined in Section 2.3.

Predictive performance was evaluated using nested cross-validation grouped by country. The outer procedure consisted of five folds and produced predictions for countries excluded from model fitting. Within each outer training sample, four-fold country-grouped cross-validation was used to select the penalty or tree hyperparameters. All preprocessing was repeated using only the observations in the relevant training fold. This included median replacement of missing control values, variable standardisation, estimation of the retained PCA transformations, and transformation of the corresponding validation observations. Information from an outer validation fold was therefore not used for preprocessing, variable screening, component construction, or hyperparameter selection. The models using the complete and reduced predictor sets were evaluated with identical outer-fold assignments so that their out-of-fold prediction errors could be compared observation by observation. The eight country-specific spline interpolations described in Section 2.4 were completed before the country-grouped folds were formed. Because the folds were defined by country, no observation from an outer validation country entered a training sample; however, each interpolated value used adjacent observations from the same country. The predictive analysis therefore assessed transfer to countries excluded from model fitting rather than forecasting their future observations.

For Ridge and Lasso, the candidate penalty values covered logarithmically spaced grids from 10^{-3} to 10^{3} and from 10^{-3} to 10^{1}, respectively. Elastic Net combined penalty values from 10^{-3} to 10^{1} with L_1-mixing parameters of 0.25, 0.50, and 0.75. For each outer training sample, the combination producing the lowest mean inner-fold RMSE was retained.

The gradient-boosting grid considered 50 or 100 trees, maximum tree depths of two or three, learning rates of 0.01 or 0.03, and minimum terminal-node sizes of three or seven. Hyperparameters were selected by minimising mean RMSE across the inner country-grouped folds. A fixed random seed of 12,345,678 was used for reproducibility.

The predictions from the five outer folds were combined to calculate out-of-fold R^2, RMSE, and mean absolute error (MAE). Performance was also reported separately for each outer fold and compared with a benchmark that assigned each validation observation the mean FDI value from the corresponding training sample. The complete nested procedure was conducted for the full predictor set and for the reduced set excluding all defence indicators, using the same country-grouped outer folds. Paired differences in mean squared prediction error were calculated between the two predictor sets and between each fitted model and its corresponding training-mean benchmark. Uncertainty intervals for these differences were obtained from 5000 bootstrap samples drawn at the country level. The bootstrap resampled countries together with their fixed out-of-fold predictions and therefore quantified uncertainty in the observed paired prediction-error

differences; it did not repeat the complete preprocessing and model-tuning procedure within every bootstrap sample. Positive full-minus-reduced differences indicated poorer performance after the defence indicators were included. Model interpretation was conditional on demonstrable out-of-country predictive performance, and the comparison was treated as evidence about the evaluated predictor sets and algorithms rather than as proof that defence information could not improve prediction under any alternative design.

Predictor-selection frequencies were calculated across the five outer training samples for Lasso and Elastic Net; a predictor retained in only one or two outer samples was not treated as stable. As the gradient-boosting model did not demonstrate reliable improvement over the benchmark or the corresponding model without defence indicators, Shapley-value attribution was not retained as substantive evidence. Table S6A–D report the overall and fold-specific performance, selected hyperparameters, and predictor-selection frequencies, while Table S6E,F report the matched comparison with the reduced predictor set and the paired country-bootstrap results. As a country-coverage sensitivity analysis, the complete nested predictive procedure was repeated after excluding Hungary, Ireland, and the Netherlands. The reduced-country predictive sample contained 335 country–year observations from 20 countries. The outer and inner folds remained grouped by country, and median replacement, variable screening, standardisation, PCA estimation, hyperparameter selection, and model fitting were all repeated within the relevant training samples. Performance was evaluated using pooled out-of-fold R^2, RMSE, and MAE, with uncertainty intervals obtained from 5000 country-level bootstrap samples. Predictor-selection frequencies were recalculated for the reduced-country sample. The results of this analysis are reported in Table S12A–E.

2.7. Panel Regressions and Temporal Predictability

Panel regressions were used to estimate conditional contemporaneous associations between each defence indicator and FDI [29]. All specifications included country fixed effects to control for time-invariant differences between countries and year fixed effects to absorb common annual shocks. The same nested sequence of control sets was applied to each defence indicator, thereby avoiding changes in the fixed-effects structure or control strategy across indicators.

$$FDI_{i,t} = \beta_0 + \beta_1 D_{i,t} + \boldsymbol{\gamma}^{\mathrm{T}} \mathbf{X}_{i,t} + \mu_i + \tau_t + \varepsilon_{i,t}$$

Here, $FDI_{i,t}$ denotes net FDI inflows as a percentage of GDP, $D_{i,t}$ is the defence indicator, and $\mathbf{X}_{i,t}$ is the relevant control vector. Country and year fixed effects are denoted by μ_i and τ_t, respectively. The coefficient β_1 therefore measures the conditional within-country association between the defence indicator and FDI after accounting for the included controls and common annual effects. The global uncertainty indices were not entered as separate main effects because their values were identical for all countries within a year and were consequently absorbed by the year fixed effects.

Military expenditure as a percentage of GDP and as a percentage of government expenditure entered the models in their original percentage scales. Military expenditure per capita and defence investment were log-transformed because of their pronounced right-skewness. GMI was entered in levels; the reciprocal-GMI specification used in the preliminary analysis was not retained because it arose from exploratory functional-form assessment. Arms imports were transformed using the inverse hyperbolic sine, which accommodates zero values without requiring a logarithmic offset. Only originally observed arms-import values were used. The common B0–B2 specifications are summarised in Table 2. As an additional measurement check, defence investment was also expressed as a percentage of GDP and as a percentage of estimated total military expenditure. Total military expenditure was calculated by multiplying GDP by military expenditure as a

percentage of GDP and dividing by 100. Both investment-intensity ratios were log-transformed and estimated using the B0–B2 specifications. These alternatives were treated as exploratory because they were constructed after the principal indicator set had been defined.

Table 2. Common panel-regression specifications.

Model	Defence Specification	Controls	Fixed Effects
B0	One defence indicator	None	Country and year
B1	One defence indicator	Log GDP per capita, trade openness, inflation, gross capital formation, and unemployment	Country and year
B2	One defence indicator	B1 controls, secondary-school enrolment, research and development expenditure, and regulatory quality	Country and year

Note: Each defence indicator was entered in a separate model. B2 was treated as the principal specification. The two primary indicators were military expenditure as a percentage of GDP and military expenditure as a percentage of government expenditure. The four secondary indicators were log military expenditure per capita, log defence investment, GMI, and inverse-hyperbolic-sine-transformed arms imports.

$$FDI_{i,t} = \beta_0 + \beta_1 D^c_{i,t} + \beta_2 \left(D^c_{i,t} \times GPR^c_t \right) + \boldsymbol{\gamma}^{\mathrm{T}} \mathbf{X}_{i,t} + \mu_i + \tau_t + \varepsilon_{i,t}$$

Moderation by global geopolitical risk was examined separately for the two primary defence indicators. The defence indicator and GPR were centred at their sample means before constructing the interaction term. The GPR main effect was absorbed by the year fixed effects, but its interaction with a country-varying defence indicator remained identifiable. Marginal defence slopes were calculated at the 10th, 50th, and 90th percentiles of GPR. The corresponding GPR values, marginal slopes, confidence intervals, and p-values under country-clustered and Driscoll–Kraay inference are reported in Table S8D. Possible changes in the primary defence–FDI associations over time were examined in additional segmented-slope models. Each primary indicator was interacted with indicators for 2014–2023 and 2022–2023, producing estimated slopes for 2005–2013, 2014–2021, and 2022–2023. The two interaction coefficients were tested jointly using country-clustered and Driscoll–Kraay covariance estimates. These models retained the B2 controls and country and year fixed effects. They were interpreted as exploratory tests of coefficient stability rather than as event-study or causal estimates, particularly because the final segment contained only two years.

Inference was assessed using three complementary procedures. First, standard errors were clustered by country, using 22 degrees of freedom in the principal 23-country sample. Second, Driscoll–Kraay standard errors were calculated to allow for heteroscedasticity, serial correlation, and general cross-sectional dependence [30]. A bandwidth of two years was used for the principal Driscoll–Kraay results, with bandwidths of one and three examined as sensitivity checks. The complete B2 results for bandwidths of one, two, and three years are reported in Table S5D. Third, restricted wild-cluster-bootstrap t-tests used 9999 replications and Webb's six-point weights [31]. Coefficients, 95% confidence intervals, and the results of all three inference procedures were reported. As these procedures address different forms of sampling uncertainty, agreement across them was treated as stronger evidence than an isolated threshold crossing under a single covariance estimator.

Residual cross-sectional dependence was assessed using Pesaran's CD test [32], with the complete results reported in Table S4. For the two primary defence indicators, Holm adjustment was applied separately within each B0–B2 specification to the p-values obtained from each inferential procedure: country-clustered, Driscoll–Kraay, and wild-cluster-bootstrap inference [33]. This allowed the multiplicity-adjusted conclusions to be

compared without treating one covariance estimator as the sole basis for inference. For the four secondary indicators, false-discovery-rate-adjusted p-values were calculated using the Benjamini–Hochberg procedure [34]. The same procedure was applied to the two exploratory defence-investment-intensity indicators within each specification. The post-2014 and post-2022 interaction coefficients were additionally assessed using a joint test for each primary defence indicator. Unadjusted and adjusted values were both reported. Moderation, segmented-slope, alternative defence-investment, and country-profile analyses were classified as exploratory.

Country-profile differences were assessed in pooled models containing interactions between the defence indicator and the profile-membership indicators. These models retained the B2 controls and both country and year fixed effects. The two interaction coefficients were tested jointly using country-clustered covariance estimates and a wild-cluster bootstrap test. This approach tested whether the slopes differed across profiles without relying on comparisons of statistical significance from separately estimated small-sample regressions. The analysis was conducted for the two primary defence indicators and GMI and was interpreted descriptively.

To express the magnitude of the primary associations on a comparable scale, the within-country standard deviations of each defence indicator and FDI were calculated after subtracting the corresponding country means in each estimation sample. The estimated B2 coefficient was multiplied by the within-country standard deviation of the defence indicator to obtain the change in FDI percentage points associated with a one-standard-deviation within-country change in that indicator. A fully standardised coefficient was additionally obtained by dividing this quantity by the within-country standard deviation of FDI. The same calculations were applied to the principal country- and period-sensitivity samples.

Dynamic sensitivity was first examined by adding one-year-lagged FDI to each primary B2 fixed-effects model while retaining the country and year effects and the three inference procedures. Because lagged dependent-variable fixed-effects estimates may be biased in panels with a limited time dimension, these models were treated as sensitivity analyses rather than preferred estimators. A parsimonious one-step difference-GMM analysis was also conducted after excluding Hungary, Ireland, and the Netherlands [35]. Lagged FDI and the defence indicator were treated as predetermined and instrumented using collapsed level instruments dated (t-2) and (t-3). The B2 controls entered in first differences, and indicators for the post-2014 and post-2022 periods were included to limit the number of time instruments. The resulting specifications used 14 instruments for 20 countries, thereby keeping the instrument count below the number of cross-sectional units [36]. Instrument validity and residual dependence were assessed using the Hansen overidentification test and the Arellano–Bond AR(1) and AR(2) tests. All dynamic results were reported irrespective of whether the diagnostic tests supported the specification.

Lagged predictability between FDI and each defence indicator was examined in both directions using the Dumitrescu–Hurlin heterogeneous-panel non-causality test [37,38]. Before estimation, each series was demeaned by year to reduce the influence of shocks shared by all countries. The direction from defence to FDI was represented by the following model:

$$FDI_{i,t} = \alpha_i + \sum_{k=1}^{K} \gamma_{i,k} FDI_{i,t-k} + \sum_{k=1}^{K} \beta_{i,k} D_{i,t-k} + \varepsilon_{i,t}.$$

The variables exchanged roles in the reverse-direction tests. The null hypothesis,

$$H_0: \beta_{i,1} = \cdots = \beta_{i,K} = 0, \qquad \forall i$$

states that past defence values do not improve the prediction of FDI in any country beyond the information contained in past FDI. Rejection of the null indicates that past values

of the predictor contain additional information for forecasting the outcome in at least a subset of countries. It does not identify those countries or establish a structural causal effect.

Common lag orders of one and two years were examined. FDI, both primary military-expenditure indicators, log defence investment, GMI, and inverse-hyperbolic-sine-transformed arms imports entered the tests in levels, while log military expenditure per capita was first-differenced. The level specification was supported by the CIPS results for the complete panel and, where relevant, the smaller balanced samples. In the 18-country defence-investment sample, the unit-root null for FDI was rejected at the 1% level. In the 13-country arms-import sample, however, it was rejected only at the 10% level; the temporal results involving arms imports were therefore treated as exploratory. Analyses of variables with incomplete observations used the largest balanced country subset containing only originally observed values. This resulted in 18 countries for defence investment and 13 countries for arms imports; missing arms-import observations were not replaced by zeros.

The conventional finite-T standardised statistic was calculated for each direction–lag combination. Because the test may be affected by dependence across countries, residual cross-sectional dependence was examined using Pesaran's CD statistic calculated from the unrestricted-model residuals. These diagnostic residuals were obtained from the alternative specification containing both the outcome's own lags and the predictor lags; they were used only to assess residual dependence and did not serve as bootstrap innovations. The conventional test results were retained as descriptive reference estimates, but the principal temporal-predictability inference was based on a restricted-residual circular moving-block bootstrap. For each country, the outcome was first estimated under the null model containing only its own lags. The resulting restricted-model residuals were centred within country, arranged by year, and used as the bootstrap innovations. Common circular blocks of annual residual vectors were resampled synchronously across countries, thereby preserving the remaining contemporaneous cross-country dependence while retaining short-range temporal dependence. Bootstrap outcome series were generated recursively under the null hypothesis, and the mean country-specific Wald statistic was recalculated for each bootstrap sample. Cross-sectional demeaning was not repeated within the bootstrap replications: it was applied once to the observed series before model estimation, while the synchronous resampling of the restricted-residual vectors was used to preserve the dependence remaining after that transformation. The principal procedure used blocks of three years and 999 bootstrap replications; block lengths of two and four years, each with 399 replications, were examined as sensitivity checks. A fixed random seed of 12,345,678 was used for reproducibility.

The Benjamini–Hochberg procedure was applied jointly to the 24 bootstrap p-values representing all combinations of defence indicator, direction, and lag order. Temporal-predictability evidence was treated as notable only when the Benjamini–Hochberg-adjusted block-bootstrap q-value was below 0.05. The conventional standardised-test p-values and their adjustments were reported for comparison but were not used as the final basis for claims when they conflicted with the dependence-robust bootstrap results. Bidirectional predictability, if supported, would indicate that the two series contained information about each other's subsequent values; it would not establish simultaneous causation or structural feedback. The original set of time-demeaned tests was also repeated after excluding Hungary, Ireland, and the Netherlands, using the same transformations, lag orders, balanced-subpanel requirements, and directional hypotheses. This reduced-country analysis was retained as a descriptive country-composition sensitivity check. The complete conventional and block-bootstrap results are reported in Table S7A,B, while the reduced-country results are reported in Table S13.

All data preparation, statistical analyses, predictive modelling, and figure generation were performed using Python 3.12.14 (Python Software Foundation, Wilmington, DE, USA). The tested computational environment used NumPy 2.3.5, pandas 2.2.3, SciPy 1.17.0, scikit-learn 1.8.0, openpyxl 3.1.5, and Matplotlib 3.10.6. The complete software environment, package versions, fixed random seeds, and execution instructions are documented in Supplementary Files.

3. Results

3.1. Data Characteristics and Preliminary Relationships

The principal panel contained 437 observations for 23 European Union countries over 2005–2023. Defence-investment analyses used 423 observations because 14 values remained unavailable. Arms-import analyses used 397 observations because the 40 originally missing values were retained as missing. The main outcome and both primary defence indicators were available for all 437 country–year observations, although the inclusion of secondary-school enrolment reduced the B2 estimation sample to 436 observations.

Table 3 shows substantial dispersion and asymmetry in several variables. FDI has a strongly right-skewed and heavy-tailed distribution, while its negative minimum represents country–years in which net disinvestment exceeded new inward investment. Defence investment and arms imports are also markedly right-skewed, reflecting a relatively small number of large observations. The three uncertainty indices use different numerical scales, so their absolute dispersion is not directly comparable.

The largest FDI value was recorded for Hungary in 2020 (105.64% of GDP), while the lowest was recorded for Hungary in 2018 (−40.11%). Other prominent positive observations included the Netherlands in 2007 (85.98%) and Ireland in 2015 (74.82%). Studentised residuals, leverage values, and Cook's distances identified these and several other observations, particularly those from Hungary and the Netherlands, as potentially influential. The observations were retained in the principal sample because they could not be classified as data errors. Their influence was examined through leave-one-country-out estimation and through a sensitivity analysis that jointly excluded Hungary, Ireland, and the Netherlands.

Table 3. Pooled descriptive statistics for FDI, defence, and global uncertainty indicators.

Variable	N	Mean	Median	SD	IQR	Minimum	Maximum
FDI net inflows, % of GDP	437	5.51	3.18	12.13	3.87	−40.11	105.64
Military expenditure, % of GDP	437	1.46	1.38	0.54	0.62	0.21	4.00
Military expenditure, % of government expenditure	437	3.28	3.01	1.29	1.52	0.89	8.74
Military expenditure per capita, EUR	437	321.45	284.31	185.57	251.11	69.43	1265.47
Defence investment, EUR million	423	1531.62	398.90	2645.25	1428.39	2.41	14,200.00
Global Militarisation Index	437	114.58	111.30	31.93	37.45	43.51	215.57
Arms imports, SIPRI TIV	397	119.26	47.00	189.29	134.00	0.00	1665.00
Geopolitical Risk Index	437	96.99	93.74	18.24	21.00	77.30	157.58
Global Economic Policy Uncertainty Index	437	156.44	148.31	65.04	94.30	63.23	306.53
World Uncertainty Index	437	21,864.82	20,001.84	7131.83	9006.09	10,173.53	40,648.58

Note: Statistics are pooled across country–year observations. Defence-investment statistics use the 423 available observations, and arms-import statistics use 397 originally observed values. Arms imports are expressed as SIPRI trend-indicator values (TIV). The three global uncertainty indices vary annually, with the same value assigned to every country in a given year.

Figure 1 complements the pooled descriptive statistics by showing the annual medians and interquartile ranges of FDI and the two primary defence indicators. Median FDI was comparatively high during the first years of the study period and subsequently fluctuated at lower levels, while its interquartile range varied substantially across years. Both military-expenditure measures declined until approximately 2014–2015 and then generally increased, although their annual distributions continued to overlap across the marked periods. The figure therefore indicates gradual changes and considerable cross-country dispersion rather than a clearly isolated shift in 2014, 2020, or 2022.

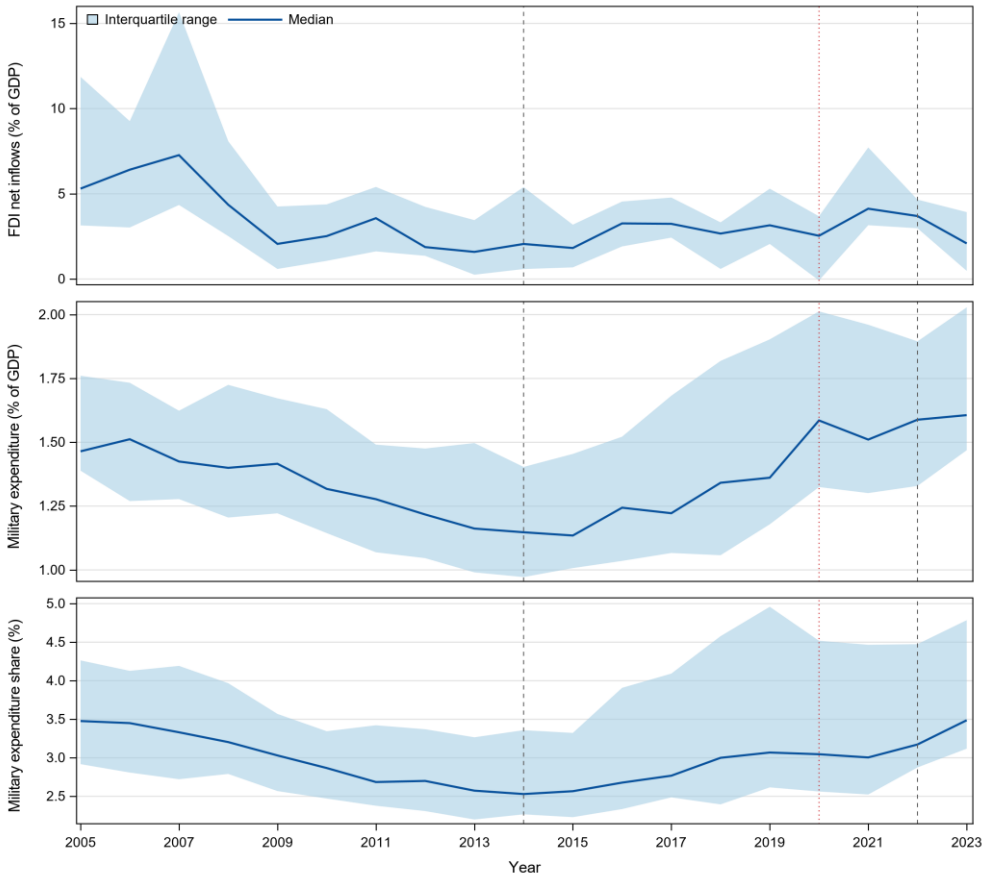


Figure 1. Annual distributions of FDI, military expenditure as a percentage of GDP, and military expenditure as a percentage of government expenditure, 2005–2023. The solid lines show annual country medians, and the shaded areas show interquartile ranges. The vertical markers identify 2014, the beginning of the main COVID-19 period in 2020, and 2022. These dates are shown for descriptive orientation and do not represent estimated causal break points.

The pooled correlations in Table 4 were small in absolute magnitude. Before adjustment for multiple testing, FDI was negatively correlated with military expenditure as a percentage of GDP, defence investment, GMI, the Global Economic Policy Uncertainty Index, and the World Uncertainty Index at the 5% level. After Benjamini–Hochberg adjustment across the nine correlations, only the association with the Global Economic Policy Uncertainty Index remained below the 5% false-discovery-rate threshold. These pooled coefficients do not account for country or year effects and should not be interpreted as conditional within-country relationships.

Table 4. Pooled Pearson correlations of FDI and the defence and global uncertainty indicators.

Indicator	N	Correlation with FDI	Unadjusted *p*-Value	BH-Adjusted *q*-Value

Military expenditure, % of GDP	437	−0.100	0.036	0.087
Military expenditure, % of government expenditure	437	−0.054	0.257	0.257
Military expenditure per capita	437	−0.084	0.078	0.118
Defence investment	423	−0.119	0.014	0.065
Global Militarisation Index	437	−0.098	0.040	0.087
Arms imports	397	−0.066	0.192	0.216
Geopolitical Risk Index	437	−0.068	0.156	0.200
Global Economic Policy Uncertainty Index	437	−0.150	0.002	0.015
World Uncertainty Index	437	−0.094	0.049	0.087

Note: Coefficients are pooled across country–year observations and do not control for country or year effects. Pairwise available observations were used. The Benjamini–Hochberg adjustment was applied jointly to the nine reported correlations. The global uncertainty indices vary only by year.

Several control variables exhibited moderate or strong pooled pairwise dependencies. This finding supported the use of penalisation and fold-specific PCA in the supplementary predictive analysis. In the principal panel regressions, multicollinearity was limited by using the eight observed controls defined for the B2 specification rather than entering the complete initial control set simultaneously. Pooled correlations alone do not establish conditional, within-country, or nonlinear predictive relationships.

In the complete 23-country sample, the CIPS statistic was −2.720 for FDI, rejecting the unit-root null at the 1% level. Rejection was obtained at the 5% level for military expenditure as a percentage of GDP (−2.339), military expenditure as a percentage of government expenditure (−2.358), and GMI (−2.367), and at the 1% level for log defence investment (−2.600) and inverse-hyperbolic-sine-transformed arms imports (−3.199). For log military expenditure per capita, the statistic of −2.197 crossed the 10% but not the 5% critical value. This variable was therefore first-differenced in the temporal-predictability analysis.

The results were generally similar after Hungary, Ireland, and the Netherlands were excluded. The unit-root null was rejected at the 1% level for FDI (−3.272), both primary military-expenditure measures (−2.596 and −2.438), log defence investment (−2.464), and transformed arms imports (−3.089), and at the 5% level for GMI (−2.290). In contrast, the reduced-sample CIPS statistic for log military expenditure per capita was −2.030 and did not reject the unit-root null even at the 10% level. The country-specific ADF and KPSS diagnostics showed substantial heterogeneity and, for several variables, differing conclusions across countries. The panel-level CIPS rejections were therefore interpreted as evidence about the joint panel rather than as confirmation that every national series was stationary. Table S1A–C report the CIPS and country-specific stationarity diagnostics, while Table S1D reports the first-difference sensitivity model.

3.2. Principal Components and Country Profiles

PCA was conducted separately for the economic, infrastructure, political-institutional, and human-capital variable groups. Table 5 reports the sampling-adequacy measures, the variance explained by PC1, and the number of components supported by parallel analysis. The infrastructure and political-institutional groups had the strongest common correlation structures. Sampling adequacy was weaker for the economic group, while the KMO value of 0.480 for human capital was below the conventional threshold of 0.50. Results based on the latter component therefore require particular caution.

Table 5. PCA diagnostics and parallel-analysis results.

Variable Group	Number of Retained Variables	Overall KMO	Variance Explained by PC1, %	Components Supported by Parallel Analysis
Economic conditions	9	0.559	25.1	3

Infrastructure	6	0.786	65.4	1
Political-institutional conditions	13	0.887	51.1	2
Human capital	4	0.480	36.0	1

Note: The number of components was determined by comparing the observed eigenvalues with the 95th percentiles of eigenvalues obtained from 1000 independently permuted datasets. PC1 was retained solely to provide one descriptive country-profile measure for each group.

PC1 explained 65.4% of the standardised variance in infrastructure and 51.1% in the political-institutional group, compared with 25.1% in the economic group and 36.0% in human capital. Parallel analysis supported one component for infrastructure and human capital, three components for the economic group, and two for the political-institutional group. A single component therefore provided only a partial representation of the economic and political-institutional variables. PC1 was nevertheless retained as a common descriptive summary for country profiling, which required one score for each of the four domains. The PCA scores were not used as controls in the principal B2 regressions. The complete eigenvalues and PC1 loading patterns are reported in Table S2A and S2B, respectively.

The four country-level PC1 means were standardised across the 23 countries before clustering. Among the K-means candidates, the three-cluster solution had the highest silhouette coefficient (0.428) and was retained. Its Davies–Bouldin index was 0.781 and its Calinski–Harabasz statistic was 20.366. Ward hierarchical clustering produced the same partition, indicating that the grouping was not specific to K-means. Table 6 presents cluster membership and centres calculated from the standardised country-level matrix. The country-specific values forming this matrix are provided in Supplementary Data S1, while Table S3 reports the validation results for all clustering methods and candidate cluster numbers.

Table 6. Country profiles and cluster centres in the standardised country-level PCA matrix.

Profile	Countries	Economic Centre	Infrastructure Centre	Political-Institutional Centre	Human-Capital Centre
Profile 1	Germany, Spain, France, Italy	−0.339	1.957	0.408	−0.834
Profile 2	Bulgaria, Czechia, Greece, Croatia, Hungary, Lithuania, Latvia, Poland, Portugal, Romania, Slovakia, Slovenia	−0.564	−0.522	−0.746	−0.398
Profile 3	Austria, Belgium, Estonia, Finland, Ireland, the Netherlands, Sweden	1.160	−0.223	1.046	1.159

Note: The centres are means of the standardised country-level PC1 averages used in K-means clustering. They describe positions relative to the 23-country sample and should be interpreted together with the component loadings in Table S2B. Because PCA orientation is arbitrary, positive values do not inherently indicate more favourable conditions. Country-specific scores are provided in Supplementary Data S1.

Profile 1 comprised Germany, Spain, France, and Italy and was distinguished primarily by its high infrastructure score. Its mean human-capital score was below the 23-country average. Profile 2 contained 12 countries and had below-average centres on all four dimensions, with the largest relative difference in the political-institutional score. Profile 3 contained seven countries and had the highest economic, political-institutional, and human-capital centres, while its infrastructure centre was slightly below the sample average. Figure S1 presents the standardised country-level scores underlying these cluster centres. It also shows variation among countries assigned to the same profile, confirming

that the clusters represent overall multivariate proximity rather than uniform national characteristics.

The profiles represent long-term similarities derived from country averages over 2005–2023. They should not be interpreted as predetermined institutional regimes, annual transitions, or causal country classifications. Subsequent heterogeneity analysis used pooled interaction models to test jointly whether the coefficients of the two primary defence indicators or GMI differed across the three profiles.

3.3. Nested Predictive-Modelling Results

The nested predictive analysis used 385 country–year observations with complete values for the six defence indicators and three global uncertainty measures. Table 7 compares the pooled out-of-fold performance obtained using the complete predictor set with the performance of otherwise equivalent models from which all six defence indicators were removed. Identical outer country-grouped folds were used for the two predictor sets. None of the models using the complete predictor set demonstrated reliable improvement over the corresponding training-mean benchmark. Elastic Net produced the highest out-of-fold R^2 among these models, but its value was only 0.019. Ridge and Lasso each produced an out-of-fold R^2 of approximately 0.011, whereas gradient boosting produced a negative value of −0.025.

Table 7. Out-of-fold predictive performance from nested country-grouped cross-validation.

Predictor Set	Method	N	Out-of-Fold R^2	RMSE	MAE
Benchmark	Training-mean	385	−0.002	12.684	6.513
Complete	Ridge	385	0.011	12.603	6.210
Complete	Lasso	385	0.011	12.605	6.210
Complete	Elastic Net	385	0.019	12.556	6.137
Complete	Gradient boosting	385	−0.025	12.835	6.611
Without defence indicators	Ridge	385	0.025	12.513	6.120
Without defence indicators	Lasso	385	0.009	12.614	6.253
Without defence indicators	Elastic Net	385	0.028	12.494	6.095
Without defence indicators	Gradient boosting	385	0.018	12.557	6.031

Note: Predictions were obtained from the outer validation folds of nested cross-validation grouped by country. The complete predictor set included all six defence indicators, whereas the reduced predictor set excluded them while retaining the three uncertainty measures and four fold-specific PCA control scores. The training-mean benchmark assigned each validation observation the mean FDI value from the corresponding outer training sample. RMSE denotes root mean squared error, and MAE denotes mean absolute error; both are expressed in FDI percentage points. Complete and reduced predictor sets used identical outer-fold assignments. Fold-specific results are reported in Table S6A,B.

The comparison with the training-mean benchmark remained uncertain. Relative to the benchmark, the complete Ridge, Lasso, and Elastic Net models reduced RMSE by 0.081, 0.079, and 0.128 percentage points, respectively, whereas gradient boosting increased it by 0.151 percentage points. However, the 95% country-bootstrap intervals for all four differences included zero. Thus, none of the complete models established a reliable improvement over prediction based only on the mean FDI value in the corresponding outer training sample.

The paired comparisons differed across modelling methods. For Ridge, Elastic Net, and gradient boosting, the models without the defence indicators produced lower RMSE values than their complete counterparts. The complete-minus-reduced RMSE differences were 0.090 for Ridge, 0.062 for Elastic Net, and 0.278 for gradient boosting; their respective

95% country-bootstrap intervals were [0.007, 0.181], [0.002, 0.126], and [0.034, 0.826]. Positive differences indicate poorer performance by the models containing the defence indicators. For Lasso, the corresponding difference was −0.009, with a 95% interval of [−0.097, 0.070], and was therefore inconclusive. Thus, including the evaluated defence indicators did not improve prediction under any of the four fitted methods and was associated with poorer out-of-country performance for three of them. These findings do not rule out predictive value under alternative variable definitions or modelling designs. The complete paired comparisons are reported in Table S6E,F.

The sparse-model selection frequencies are reported in Table 8. The economic PC1 score was retained in all five outer training samples by both Lasso and Elastic Net. The Global Economic Policy Uncertainty Index was retained in four Lasso samples and all five Elastic Net samples, while the infrastructure PC1 score was retained in two and three samples, respectively. GPR was selected in all five Elastic Net samples but in only one Lasso sample. Selection frequencies and mean standardised coefficients are reported in Table S6D.

Table 8. Predictor-selection frequency across the five outer training samples.

Predictor	Lasso, %	Elastic Net, %
Military expenditure, % of GDP	0	20
Military expenditure, % of government expenditure	20	20
Log military expenditure per capita	20	40
Log defence investment	0	60
Global Militarisation Index	0	40
Inverse-hyperbolic-sine-transformed arms imports	40	20
Geopolitical Risk Index	20	100
Global Economic Policy Uncertainty Index	80	100
World Uncertainty Index	0	20
Economic PC1 score	100	100
Infrastructure PC1 score	40	60
Political-institutional PC1 score	0	20
Human-capital PC1 score	20	60

Note: Selection frequency is the percentage of the five outer training samples in which the fitted coefficient was non-zero after the penalty had been selected by inner country-grouped cross-validation. Ridge is omitted because it does not set coefficients exactly to zero. Frequencies of 20%, 40%, 60%, 80%, and 100% correspond to selection in one, two, three, four, and five outer training samples, respectively. Selection is a predictive result and does not indicate statistical significance or causality.

No defence indicator was selected consistently. Military expenditure as a percentage of GDP was not selected by Lasso and was retained in only one Elastic Net sample. Military expenditure as a percentage of government expenditure was selected in one outer training sample by each method. Log military expenditure per capita was retained in one Lasso and two Elastic Net samples, while log defence investment was not selected by Lasso but was retained in three Elastic Net samples. GMI was retained in two Elastic Net samples and none of the Lasso samples, whereas arms imports were retained in two Lasso samples and one Elastic Net sample. The selection patterns therefore do not support a stable incremental predictive contribution from any defence indicator.

As a separate country-composition sensitivity analysis, the nested predictive analysis using the complete predictor set was repeated after excluding Hungary, Ireland, and the Netherlands, reducing the sample to 335 country–year observations. Gradient boosting produced an overall out-of-fold R^2 of 0.026, an RMSE of 4.924 percentage points, and an MAE of 2.845 percentage points. Its country-bootstrap 95% confidence interval for R^2,

however, ranged from −0.256 to 0.163. Ridge produced a similarly small positive R^2 of 0.024, whereas the corresponding values for Lasso and Elastic Net were −0.016 and −0.008. The confidence intervals for all four methods included zero, predictor selection remained unstable, and no defence indicator was retained in all five outer training samples by either Lasso or Elastic Net. The complete performance estimates, fold-specific results, tuning parameters, fold-specific coefficients, and selection frequencies for this sensitivity analysis are reported in Table S12A–E.

Taken together, the predictive results indicate that the available predictors explained little transferable variation in FDI for countries excluded from model fitting. None of the complete models demonstrated reliable improvement over the training-mean benchmark. In the paired comparisons, Ridge, Elastic Net, and gradient boosting performed better without the defence indicators, while the Lasso comparison was inconclusive. Predictor retention varied across the outer training samples, and no defence indicator was selected consistently. The country-composition sensitivity analysis led to the same broader conclusion: although gradient boosting and Ridge produced small positive point estimates of out-of-fold R^2 after Hungary, Ireland, and the Netherlands were excluded, their confidence intervals were wide and included zero. These results do not support substantive interpretation of model-specific importance measures, and Shapley-value results from the preliminary analysis were therefore not retained.

3.4. Panel Regression and Robustness Results

Table 9 reports the fixed-effects panel estimates. All specifications include country and year fixed effects. Because the residuals exhibited strong cross-sectional dependence, the interpretation gives particular weight to Driscoll–Kraay inference, supplemented by country-clustered standard errors and the wild-cluster bootstrap. The estimated coefficients represent conditional contemporaneous associations and should not be interpreted as causal effects.

The two primary defence indicators produced positive coefficients in all three specifications. In the fully adjusted B2 models, a one-percentage-point increase in military expenditure relative to GDP was associated with a 5.428-percentage-point increase in FDI relative to GDP. The corresponding Driscoll–Kraay 95% confidence interval ranged from 1.377 to 9.478. Military expenditure relative to government expenditure had a B2 coefficient of 1.844, with a Driscoll–Kraay 95% confidence interval from 0.309 to 3.378. Before adjustment for the two primary hypotheses, both results were below the 5% threshold under country-clustered, Driscoll–Kraay, and wild-cluster-bootstrap inference. After Holm adjustment, the two country-clustered p-values were 0.053 and therefore narrowly exceeded 0.05. The Holm-adjusted wild-cluster-bootstrap p-values were 0.037, while the Holm-adjusted Driscoll–Kraay p-values were 0.023 when the principal bandwidth of two years was used. The Driscoll–Kraay conclusions were unchanged with bandwidths of one and three years. Thus, the positive coefficients were supported by dependence-robust and bootstrap inference after multiplicity adjustment but not by the adjusted country-clustered tests.

Because a one-percentage-point change exceeded the within-country variation observed in most countries, effect sizes were also calculated for an increase of one within-country standard deviation in each indicator. In the main B2 sample, such an increase corresponded to 1.511 FDI percentage points for military expenditure relative to GDP and 1.251 points for military expenditure relative to government expenditure. Their Driscoll–Kraay 95% confidence intervals were 0.383 to 2.638 and 0.210 to 2.292 percentage points, respectively. Figure 2 compares these standardised estimates with the corresponding estimates from the principal country and period sensitivity analyses; the complete calculations are reported in Table S10.

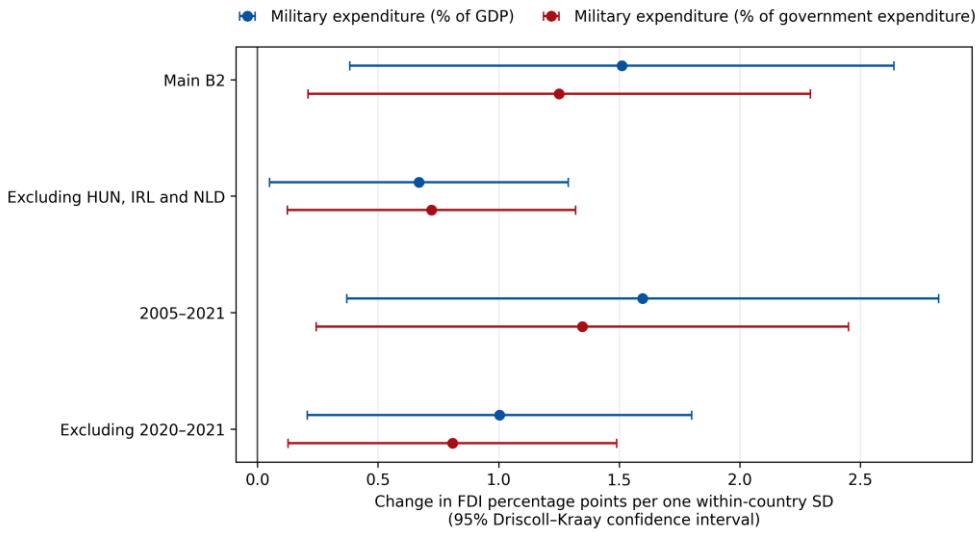


Figure 2. Within-country standardised associations between the primary military-expenditure indicators and FDI. Points show the estimated change in FDI as a percentage of GDP associated with a one within-country standard-deviation increase in the indicated defence measure. Horizontal lines show 95% Driscoll–Kraay confidence intervals. All estimates are based on the fully adjusted B2 specification with country and year fixed effects.

Among the secondary indicators, log military expenditure per capita and GMI had positive coefficients in the fully adjusted models. Their Driscoll–Kraay p-values remained below 0.05 after Benjamini–Hochberg adjustment across the four secondary indicators. Defence investment and arms imports provided no corresponding evidence. The conclusion for defence investment remained unchanged when it was rescaled to improve comparability with the outcome and the other defence indicators. In the B2 models, log defence investment as a percentage of GDP had a coefficient of 1.975 ($p_{DK} = 0.093$), while log defence investment as a percentage of estimated military expenditure had a coefficient of 1.506 ($p_{DK} = 0.186$). Both Benjamini–Hochberg-adjusted q-values were 0.186. Descriptive statistics for the constructed ratios and the complete B0–B2 results are reported in Table S9A and S9B, respectively. Because military expenditure per capita is closely related to the two primary expenditure measures and the secondary hypotheses were exploratory, the positive secondary results should not be treated as independent confirmations of the primary findings.

Table 9. Fixed-effects panel estimates for FDI net inflows as a percentage of GDP.

Indicator	Specification	N	Coefficient	Country-Clustered p	Driscoll–Kraay 95% CI	Driscoll–Kraay p	Wild-Bootstrap p	Adjusted p or q
Military expenditure, % of GDP	B0	437	6.015	0.039	[3.050, 8.980]	<0.001	0.035	0.001
Military expenditure, % of GDP	B1	437	6.813	0.029	[3.904, 9.722]	<0.001	0.023	<0.001
Military expenditure, % of GDP	B2	436	5.428	0.027	[1.377, 9.478]	0.011	0.018	0.023
Military expenditure, % of government expenditure	B0	437	1.987	0.041	[0.945, 3.028]	<0.001	0.055	0.001
Military expenditure, % of government expenditure	B1	437	2.385	0.030	[1.227, 3.544]	<0.001	0.039	<0.001
Military expenditure, % of government expenditure	B2	436	1.844	0.029	[0.309, 3.378]	0.021	0.031	0.023

Log military expenditure per capita	B2 436	8.943	0.018 [3.251, 14.635]	0.004	0.017	0.016
Log defence investment	B2 422	1.903	0.323 [−0.419, 4.225]	0.102	0.339	0.136
GMI	B2 436	0.103	0.022 [0.024, 0.182]	0.013	0.046	0.026
Inverse-hyperbolic-sine arms imports	B2 396	0.454	0.488 [−0.278, 1.187]	0.209	0.534	0.209

Note: The dependent variable is FDI net inflows as a percentage of GDP. All models include country and year fixed effects. B0 contains the defence indicator only; B1 adds the core economic controls; and B2 adds the complete prespecified control set. Country-clustered tests use 22 degrees of freedom. Driscoll–Kraay estimates use a bandwidth of two years and 18 time-series degrees of freedom. Wild-cluster-bootstrap p-values are based on 9999 restricted bootstrap replications with Webb's six-point weights. The final column reports Holm-adjusted Driscoll–Kraay p-values for the primary indicators within each specification and Benjamini–Hochberg-adjusted Driscoll–Kraay q-values for the four secondary B2 models.

Table 10 summarises the sensitivity, moderation, and country-profile analyses. Winsorising FDI at the first and ninety-ninth percentiles reduced both primary coefficients but did not reverse their signs. The military-expenditure-to-GDP result remained below 0.05 under all three inferential procedures. The government-expenditure-share result remained below 0.05 using country-clustered and Driscoll–Kraay inference, while its wild-bootstrap p-value was 0.051.

Excluding Hungary, Ireland, and the Netherlands reduced both primary coefficients substantially but did not reverse their signs. The coefficient for military expenditure relative to GDP fell to 2.330, while that for military expenditure relative to government expenditure fell to 1.019. Country-clustered and Driscoll–Kraay p-values remained below 0.05, including after Holm adjustment across the two primary hypotheses. However, the corresponding wild-cluster-bootstrap p-values were 0.075 and 0.100. On the within-country standard-deviation scale, the estimated associations fell to 0.669 and 0.722 FDI percentage points, respectively. Thus, the results remained positive but were weaker and were not supported uniformly across the three inferential procedures. Leave-one-country-out estimates were also positive, although their statistical support varied with the omitted country and inference method. Complete leave-one-country-out estimates and the observations with the largest Cook's distances are reported in Table S5A,B.

The alternative common per-capita specification produced mixed evidence. The coefficient was positive and below 0.05 under Driscoll–Kraay inference but not under country-clustered or wild-bootstrap inference. Because the CIPS results for log military expenditure per capita were weaker than those for the other indicators, an additional first-difference B2 model related the annual change in FDI to the annual change in log military expenditure per capita and the corresponding changes in the eight controls. The coefficient was 9.779, but it was not supported by country-clustered ($p = 0.365$), Driscoll–Kraay ($p = 0.126$), or wild-cluster-bootstrap inference ($p = 0.473$). The positive level-based result for military expenditure per capita was therefore not reproduced when the analysis was expressed entirely in annual changes.

In the available-case EU27 extension, neither primary coefficient was distinguishable from zero. Military expenditure relative to GDP had a coefficient of 6.269, with country-clustered, Driscoll–Kraay, and wild-cluster-bootstrap p-values of 0.185, 0.155, and 0.104, respectively. The coefficient for military expenditure relative to government expenditure was 0.484, with corresponding p-values of 0.833, 0.753, and 0.845. Re-extracting the original 23 countries from the EU27 dataset reproduced the principal 23-country estimates, confirming that the extension did not use a different construction of the common variables.

Only 14 years—2005–2011 and 2013–2019—contained complete observations for all 27 countries. Over this common period, the 23-country coefficients were 4.044 for military expenditure relative to GDP and 1.402 for military expenditure relative to government expenditure; neither was supported consistently across all three inferential procedures. In the corresponding EU27 sample, the coefficients were 7.614 and −0.057, respectively, and neither was statistically supported. The contrast between the 23-country and EU27 results therefore reflects country composition together with the shorter common observation window, rather than a change in the construction or processing of the variables. The matched-sample comparisons and country-level completeness information are reported in Table S14A,B.

The interactions between geopolitical risk and the two primary defence indicators were small and statistically unsupported. The interaction p-values ranged from 0.539 to 0.943 under country-clustered and Driscoll–Kraay inference. The pooled country-profile models likewise provided no evidence that the defence coefficients differed across profiles. Joint wild-bootstrap p-values were 0.924 for military expenditure relative to GDP, 0.963 for military expenditure relative to government expenditure, and 0.189 for GMI.

Table 10. Sensitivity, moderation, and country-profile results.

Analysis	Indicator or Test	Estimate or Test Statistic	Country-Clustered p	Driscoll–Kraay p	Wild-Bootstrap p
FDI winsorised at 1st and 99th percentiles	Military expenditure, % of GDP	4.385	0.016	0.007	0.019
FDI winsorised at 1st and 99th percentiles	Military expenditure, % of government expenditure	1.430	0.024	0.014	0.051
Hungary, Ireland, and the Netherlands excluded	Military expenditure, % of GDP	2.330	0.026	0.036	0.075
Hungary, Ireland, and the Netherlands excluded	Military expenditure, % of government expenditure	1.019	0.016	0.021	0.100
Common per-capita scaling	Log military expenditure per capita	3.007	0.107	0.026	0.144
First-difference B2 model	Change in log military expenditure per capita	9.779	0.365	0.126	0.473
EU27 available-case extension	Military expenditure, % of GDP	6.269	0.185	0.155	0.104
EU27 available-case extension	Military expenditure, % of government expenditure	0.484	0.833	0.753	0.845
Geopolitical-risk moderation	GDP-share interaction	0.003	0.895	0.943	—
Geopolitical-risk moderation	Government-expenditure-share interaction	0.007	0.539	0.754	—
Country-profile differences	GDP-share interactions, joint test	$\chi^2(2) = 0.148$	0.929	—	0.924
Country-profile differences	Government-expenditure-share interactions, joint test	$\chi^2(2) = 0.089$	0.957	—	0.963
Country-profile differences	GMI interactions, joint test	$\chi^2(2) = 3.910$	0.142	—	0.189

Note: The sensitivity models retain the B2 controls and country and year fixed effects. The common-scaling model uses the inverse-hyperbolic-sine transformation of FDI per capita and log military expenditure per capita; its coefficient is therefore not directly comparable with the coefficients from the percentage-of-GDP models. The first-difference model relates annual changes in FDI to annual changes in log military expenditure per capita and the corresponding first differences of the B2 controls. Its coefficient is not directly comparable with either the level-based per-capita estimate or the percentage-of-GDP estimates. The EU27 models use available country–year observations. Moderation estimates are coefficients of the interactions between the mean-centred Geopolitical Risk Index and the indicated defence measure. Country-profile rows report joint tests of the two defence-

indicator-by-profile interaction terms. An em dash indicates that the corresponding inference method was not applied.

Additional period sensitivity analyses are reported in Table S8A. Restricting the sample to 2005–2021 retained positive coefficients for both primary indicators. Both were below the 5% threshold under Driscoll–Kraay and wild-cluster-bootstrap inference, although country-clustered inference was weaker for military expenditure relative to GDP. Excluding the COVID-19 years 2020–2021 also retained both positive coefficients under Driscoll–Kraay inference, but the government-expenditure-share estimate was not supported by country-clustered or wild-cluster-bootstrap inference. Models allowing the defence–FDI slopes to change after 2014 and again after 2022 provided no joint evidence of structural change for either primary indicator. The country-clustered joint p-values were 0.663 and 0.301, and the corresponding Driscoll–Kraay values were 0.335 and 0.111. Period-specific slopes, particularly those for 2022–2023, are reported descriptively in Table S8B,C because the final period contains only two years.

Dynamic sensitivity models added one lag of FDI to the fully adjusted fixed-effects specifications. The defence coefficients remained positive at 4.762 for military expenditure relative to GDP and 1.589 for military expenditure relative to government expenditure. Their Holm-adjusted Driscoll–Kraay p-values were 0.015, whereas the Holm-adjusted country-clustered p-values were 0.058. The unadjusted wild-cluster-bootstrap p-values were 0.010 and 0.016. These estimates are treated as sensitivity results because lagged-outcome fixed-effects estimators may be biased in panels with a limited time dimension. Parsimonious difference-GMM models did not reject second-order serial-correlation tests and used 14 instruments for 20 countries, but the Hansen tests rejected both instrument sets. The difference-GMM estimates were therefore considered diagnostically unreliable and were not used as confirming evidence. Complete dynamic-model results are reported in Table S11A,B.

Overall, the panel analysis provides evidence of positive contemporaneous associations between FDI and the two military-expenditure measures in the main 23-country sample. In the fully adjusted models, both unadjusted results were below 0.05 under country-clustered, Driscoll–Kraay, and wild-cluster-bootstrap inference. After adjustment for the two primary hypotheses, support remained under Driscoll–Kraay and wild-cluster-bootstrap inference but not under country-clustered inference. The estimated magnitudes declined when Hungary, Ireland, and the Netherlands were excluded, and neither primary association was supported in the available-case EU27 extension. Matched-period comparisons indicate that the EU27 divergence reflects country composition and the shorter complete observation window rather than differences in variable construction. The common per-capita and first-difference specifications also produced weaker evidence. No evidence was obtained that the coefficients changed jointly after 2014 and 2022 or differed systematically across geopolitical-risk levels or country profiles. Lagged-outcome fixed-effects models retained positive coefficients, but the difference-GMM diagnostics were unsatisfactory. The panel findings should therefore be interpreted as sample- and specification-dependent contemporaneous associations rather than as general, dynamic, or causal effects of defence expenditure on FDI.

3.5. Temporal Predictability Between Defence Indicators and FDI

Dumitrescu–Hurlin tests were used to examine whether lagged values of each defence indicator contained incremental information for predicting FDI and whether predictability was also observed in the reverse direction. Each series was cross-sectionally demeaned by year before testing, and common lag orders of one and two years were examined. The conventional standardised-test results are reported in Table S7A. Residual

diagnostics nevertheless indicated that cross-sectional demeaning did not eliminate dependence across countries. Pesaran's CD test applied to the unrestricted-model residuals rejected residual cross-sectional independence in 16 of the 24 direction–lag specifications. The principal inference was therefore based on the restricted-residual circular moving-block bootstrap described in Section 2.7.

Table 11 reports the results obtained using the principal three-year block length. Four of the 24 unadjusted bootstrap p-values were below 0.05. These concerned FDI predicting military expenditure relative to GDP at one and two lags, GMI predicting FDI at one lag, and FDI predicting GMI at one lag. The bootstrap p-value for FDI predicting arms imports at one lag was exactly 0.050 and was therefore not classified as being below 0.05. None of the 24 results remained below the 5% false-discovery-rate threshold after Benjamini–Hochberg adjustment; the smallest adjusted q-value was 0.165. Accordingly, the dependence-robust analysis did not establish multiplicity-adjusted temporal predictability in either direction for any of the six defence indicators.

Table 11. Dependence-robust bootstrap inference for panel temporal predictability.

Direction of Predictability	Countries	One Lag: Bootstrap p (q)	Two Lags: Bootstrap p (q)
Military expenditure, % of GDP → FDI	23	0.097 (0.214)	0.462 (0.554)
FDI → Military expenditure, % of GDP	23	0.041 (0.165)	0.038 (0.165)
Military expenditure, % of government expenditure → FDI	23	0.063 (0.168)	0.386 (0.519)
FDI → Military expenditure, % of government expenditure	23	0.052 (0.165)	0.055 (0.165)
Δ Log military expenditure per capita → FDI	23	0.298 (0.447)	0.503 (0.575)
FDI → Δ Log military expenditure per capita	23	0.638 (0.647)	0.552 (0.602)
Log defence investment → FDI	18	0.410 (0.519)	0.181 (0.310)
FDI → Log defence investment	18	0.224 (0.358)	0.101 (0.214)
GMI → FDI	23	0.011 (0.165)	0.055 (0.165)
FDI → GMI	23	0.035 (0.165)	0.178 (0.310)
Inverse-hyperbolic-sine-transformed arms imports → FDI	13	0.411 (0.519)	0.647 (0.647)
FDI → Inverse-hyperbolic-sine-transformed arms imports	13	0.050 (0.165)	0.107 (0.214)

Note: Each cell reports the restricted-residual circular moving-block-bootstrap p-value followed by the Benjamini–Hochberg-adjusted q-value in parentheses. The adjustment was applied jointly to all 24 direction–lag tests. The principal procedure used a block length of three years and 999 bootstrap replications. Restricted-model residuals were used as the bootstrap innovations, while unrestricted-model residuals were used only for the Pesaran CD diagnostic. Annual residual vectors were resampled synchronously across countries to preserve the remaining contemporaneous cross-country dependence. The null hypothesis is that the variable preceding the arrow does not improve prediction of the variable following the arrow in any country in the panel. The symbol Δ denotes the first annual difference. Defence-investment tests used the 18-country balanced panel, while arms-import tests used the 13 countries with complete originally observed series. Missing arms-import observations were not replaced with zeros.

The multiplicity-adjusted conclusion was not sensitive to moderate changes in block length. Block lengths of two and four years yielded three and six unadjusted bootstrap p-values below 0.05, respectively, compared with four under the principal three-year specification. However, no direction–lag combination remained below the 5% false-discovery-rate threshold after Benjamini–Hochberg adjustment under any of the three block lengths; the smallest adjusted q-value was 0.165 in each case. Isolated unadjusted results therefore occurred under every block specification, whereas the multiplicity-adjusted conclusion remained unchanged.

The contrast with the conventional Dumitrescu–Hurlin inference is substantively important. The conventional standardised statistics suggested several bidirectional or one-directional relationships after Benjamini–Hochberg adjustment. Once residual cross-country dependence and short-range temporal dependence were incorporated into the resampling procedure, these relationships were no longer supported after adjustment. The conventional tests are retained in Table S7A for transparency, while the complete bootstrap results, residual CD statistics, and block-length sensitivity estimates are reported in Table S7B.

The conventional time-demeaned analysis was also repeated after Hungary, Ireland, and the Netherlands were excluded. Some rejections based on the adjusted conventional standardised-test p-values persisted in that reduced sample, as reported in Table S13. However, because these reduced-country results were not confirmed by the dependence-robust block-bootstrap evidence in the principal sample, they are interpreted only as descriptive indications of country-composition sensitivity.

Overall, the temporal analysis does not provide multiplicity-adjusted, dependence-robust evidence that lagged defence indicators predict FDI or that lagged FDI predicts the defence indicators. This conclusion does not demonstrate the absence of every possible temporal relationship. The annual series are short, the balanced panels for defence investment and arms imports are smaller, and the tests evaluate only common one- and two-year lag structures. In addition, the unit-root null for FDI in the 13-country arms-import sample was rejected only at the 10% level, so the corresponding temporal results should be regarded as exploratory. The results nevertheless show that the temporal relationships suggested by the conventional test statistics are not sufficiently robust to support substantive claims once cross-country dependence and multiple testing are addressed jointly.

4. Discussion

4.1. Interpretation of the Main Findings

The findings do not support a single relationship between defence-sector intensity and FDI. The results vary across defence indicators, model specifications, samples, inferential procedures, and statistical questions. The clearest evidence concerns positive contemporaneous associations between FDI and the two primary military-expenditure measures in the main 23-country panel. However, these associations weakened under several sensitivity analyses, were not reproduced in the EU27 extension, and were not accompanied by reliable out-of-country predictive performance. The evidence is therefore conditional rather than consistent with a general effect of defence-sector activity on foreign investment.

Regarding RQ1, military expenditure relative to GDP and military expenditure relative to government expenditure had positive coefficients across the B0–B2 specifications. In the fully adjusted B2 models, both unadjusted results were below the 5% threshold under country-clustered, Driscoll–Kraay, and wild-cluster-bootstrap inference. After Holm adjustment for the two primary hypotheses, support remained under Driscoll–Kraay and wild-cluster-bootstrap inference but not under country-clustered inference, for which both adjusted p-values were 0.053. The conclusions were not invariant to the sensitivity design. Winsorising FDI preserved the positive results, whereas excluding Hungary, Ireland, and the Netherlands reduced the estimated magnitudes and produced wild-cluster-bootstrap p-values above 0.05. The common per-capita specification yielded mixed inferential results, and the first-difference model for military expenditure per capita was unsupported. Neither primary association was supported in the available-case EU27 extension. Matched-period comparisons indicate that the EU27 divergence reflects country composition together with the shorter complete observation period rather than differences in variable construction.

The secondary indicators reinforce the importance of measurement. Log military expenditure per capita and GMI had positive coefficients in the fully adjusted level models, and their Benjamini–Hochberg-adjusted Driscoll–Kraay q-values were below 0.05. Defence investment and arms imports did not provide corresponding evidence. One possible explanation is that aggregate expenditure measures capture relatively persistent fiscal or security commitments, whereas defence investment and arms transfers may be more strongly affected by irregular procurement cycles, implementation delays, individual transactions, and reporting differences. These mechanisms were not observed or tested directly and should not be treated as findings. The positive coefficients likewise do not establish that an increase in military expenditure would itself attract FDI.

Regarding RQ2, the predictive analyses provided little evidence that the defence indicators contributed stable information for countries excluded from model fitting. None of the complete models demonstrated reliable improvement over the corresponding training-mean benchmark. In the paired comparisons, removing the defence indicators improved the out-of-country RMSE for Ridge, Elastic Net, and gradient boosting, whereas the Lasso comparison was inconclusive. Predictor retention also varied across the outer training samples. Some defence indicators entered individual Lasso or Elastic Net models, but none was selected consistently across all five samples. The same broader conclusion was obtained after Hungary, Ireland, and the Netherlands were excluded: the small positive out-of-fold R^2 estimates for Ridge and gradient boosting were accompanied by wide confidence intervals that included zero. These results mean that the positive panel coefficients should not be interpreted as evidence that defence expenditure enables reliable prediction of FDI in a country excluded from model fitting.

The analyses addressing RQ3 did not identify systematic heterogeneity. The interactions between geopolitical risk and the two primary military-expenditure measures were small and statistically unsupported. Formal joint tests also provided no evidence that the coefficients for the primary indicators or GMI differed across the three country profiles. These results do not demonstrate that security conditions or national characteristics are irrelevant. The Geopolitical Risk Index assigned the same annual value to every country, while the profiles represented broad groupings based on long-term economic, infrastructural, institutional, and human-capital characteristics. More detailed measures of national security exposure or defence-industrial structure might reveal variation that the indicators used here could not capture.

Regarding RQ4, the temporal-predictability analysis produced four unadjusted bootstrap p-values below 0.05 under the principal three-year block specification. These concerned FDI predicting military expenditure relative to GDP at one and two lags, GMI predicting FDI at one lag, and FDI predicting GMI at one lag. However, none of the 24 direction–lag tests remained below the 5% false-discovery-rate threshold after Benjamini–Hochberg adjustment. The same multiplicity-adjusted conclusion was obtained using block lengths of two and four years. The dependence-robust analysis therefore did not establish temporal predictability in either direction for any defence indicator. The arms-import results warrant particular caution because they were based on only 13 countries and the corresponding sample-specific stationarity evidence for FDI was weaker. More generally, these tests do not estimate the sign or magnitude of an effect, identify the countries in which predictability might occur, or demonstrate structural causality.

The findings across the different methods are not contradictory because the methods evaluate different properties of the data. Fixed-effects regressions estimate conditional contemporaneous associations, regularised regression and gradient boosting assess out-of-country predictive performance, interaction models test coefficient differences, and the temporal-predictability analysis examines whether lagged values contain incremental information. Taken together, the results indicate that selected expenditure-based measures

of defence intensity are associated with FDI in the main contemporaneous panel specifications, but these associations do not translate into reliable out-of-country prediction or multiplicity-adjusted, dependence-robust temporal predictability. They also do not generalise uniformly across samples, variable definitions, or analytical procedures. From the perspective of sustainable economic development, defence expenditure should therefore not be treated as either an automatic stimulus to international investment or an inevitable impediment to it. Its relationship with FDI depends on how defence activity and investment are measured and on the countries and statistical questions under examination.

4.2. Comparison with Previous Research

Previous research provides competing explanations for the relationship between defence expenditure and foreign investment. Defence commitments may improve perceived security, reinforce state capacity, and generate demand for infrastructure or advanced technologies. Conversely, defence expenditure may increase fiscal pressure, displace productive public investment, or signal exposure to geopolitical threats. The present findings do not identify one of these mechanisms as dominant. Instead, they show that the estimated relationship varies with the defence indicator, sample, model specification, and statistical objective.

The positive contemporaneous coefficients for military expenditure relative to GDP or government expenditure are directionally consistent with Douch [8], who reported a positive relationship between military expenditure and FDI in Middle Eastern and North African countries after accounting for political stability and macroeconomic conditions. However, that study used a dynamic panel model for countries in a different institutional and security setting, whereas the present estimates describe contemporaneous within-country associations in the EU. In the main sample, the two primary coefficients remained supported after Holm adjustment under Driscoll–Kraay and wild-cluster-bootstrap inference, but not under country-clustered inference. Their magnitudes declined when Hungary, Ireland, and the Netherlands were excluded, and neither coefficient was supported by wild-cluster-bootstrap inference in that reduced sample. The associations were also not reproduced in the EU27 extension. Possible EU-specific explanations include collective-security commitments and participation in international procurement or technology networks, but these channels were not measured or tested. The results therefore provide qualified evidence of a positive association in the main panel rather than support for a common defence-expenditure effect.

The sensitivity to country composition is also consistent with recognised limitations of aggregate FDI statistics. Damgaard et al. [39] showed that recorded FDI can be strongly affected by investment routed through special-purpose entities and offshore financial centres without a corresponding link to productive activity in the reporting economy. Ireland and the Netherlands are particularly relevant to this distinction because large multinational restructurings and pass-through flows can generate movements in headline FDI that are difficult to relate to annual changes in national defence expenditure. Excluding these countries together with Hungary reduced the estimated associations but did not remove them entirely. The results therefore concern reported net FDI inflows as a percentage of GDP and should not be interpreted automatically as relationships with new productive investment.

The results differ from Aziz and Khalid [7], who found a negative long-run relationship between military expenditure and FDI in developing countries without armed conflict and reported that conflict exposure modified this relationship. Bhandari et al. [9] likewise emphasised that conflict dynamics and institutional conditions can alter the association. Differences in country composition, security exposure, defence systems, observation periods, and econometric estimands may explain why the results do not coincide. In

particular, EU membership and collective-security arrangements may affect how defence expenditure is interpreted by investors, but the present study did not observe investor perceptions or test the underlying transmission mechanisms directly.

The indicator-specific findings further show why aggregate military expenditure cannot represent every form of defence activity. Military expenditure per capita and GMI produced positive secondary coefficients in the level-based panel models, whereas defence investment and arms imports did not. The defence-investment conclusion remained unchanged when investment was expressed as a percentage of GDP or estimated total military expenditure, suggesting that the lack of statistical support was not attributable solely to the original monetary scale. Procurement expenditure and arms transfers may be affected by irregular acquisition cycles, implementation delays, whether purchases are domestic or imported, and whether suppliers participate in national or international production networks. These characteristics were unavailable in the annual country-level data and were not tested as mediating mechanisms. They are therefore possible explanations for the indicator differences rather than findings of the study. The absence of an aggregate association for defence investment or arms imports should therefore not be interpreted as evidence that individual procurement projects or defence-industrial linkages have no economic consequences.

The weak out-of-country predictive performance also qualifies comparisons with studies applying machine-learning methods to FDI. Such methods can accommodate correlated predictors and nonlinear patterns [40], but successful in-sample fitting does not ensure transfer to countries omitted from model estimation. In the present analysis, none of the models using the complete predictor set demonstrated reliable improvement over the training-mean benchmark. Moreover, Ridge, Elastic Net, and gradient boosting produced significantly lower paired out-of-fold errors when the defence indicators were omitted, whereas the Lasso comparison was inconclusive. This does not establish that defence-related information could never improve FDI prediction, but it shows that the six indicators used here did not add transferable predictive information under the evaluated algorithms and validation design. Removing Hungary, Ireland, and the Netherlands did not alter the broader conclusion of weak and uncertain transferability.

The prominence of conventional investment conditions broadly agrees with the wider FDI literature. Islam and Beloucif [5] identified market size, trade openness, infrastructure, labour costs, macroeconomic stability, human capital, and growth prospects as recurrent FDI determinants, while Darmo et al. [41] reported the relevance of economic potential, productivity, taxation, wages, education, openness, transport infrastructure, and institutional conditions in the Visegrád countries. However, explanatory relevance in previous studies does not guarantee stable out-of-country prediction in a comparatively small panel. In the present analysis, the control dimensions provided at least as much predictive information without the defence indicators as with them. This result concerns incremental prediction under the selected models and should not be interpreted as evidence that conventional FDI determinants causally eliminate a defence-related association.

The geopolitical-risk results require a similarly cautious comparison. Caldara and Iacoviello [4] linked geopolitical-risk shocks to weaker investment and economic activity, while Truong et al. [42] found that the relationship between geopolitical risk and FDI in Vietnam differed across time horizons. The present study did not find evidence that the global Geopolitical Risk Index modified the coefficients of the two primary military-expenditure measures. This result does not contradict studies based on country-specific exposure or realised conflict. The index used here assigned the same annual value to every country and could only identify whether a common global change was associated with differences in the defence–FDI slope.

The temporal results extend the comparison beyond contemporaneous associations but provide less support than the conventional Dumitrescu–Hurlin statistics initially suggested. The standardised tests produced several adjusted rejections, including relationships involving the primary expenditure measures and GMI. However, residual dependence remained across countries, and none of the 24 direction–lag combinations remained below the 5% false-discovery-rate threshold under the dependence-robust moving-block bootstrap. This finding qualifies comparisons with studies reporting delayed defence–FDI relationships and reinforces the distinction between contemporaneous association and temporal prediction. The annual data do not establish whether defence expenditure precedes changes in FDI, whether FDI precedes changes in defence indicators, or whether both respond to common omitted developments. Possible channels involving fiscal capacity, protection of foreign-owned assets, strategic infrastructure, or security perceptions therefore remain hypotheses for future research rather than empirically identified mechanisms.

4.3. Theoretical and Policy Implications

The first theoretical implication is that defence-sector intensity cannot be represented adequately by a single interchangeable measure. Military expenditure relative to GDP, its share of government expenditure, expenditure per capita, defence investment, arms imports, and GMI describe different fiscal, budgetary, structural, and procurement-related features of national defence activity. Their empirical behaviour differed across the panel-regression, predictive, and temporal analyses. Conclusions obtained for one indicator should therefore not be extended automatically to the defence sector as a whole.

The positive coefficients for selected expenditure-based measures are compatible with several competing explanations. Sustained defence commitments may coincide with deterrence, institutional capacity, technological demand, or participation in international security and production networks. FDI and defence expenditure may instead respond jointly to omitted economic or geopolitical developments, while greater expenditure may also create fiscal pressure or displace other productive allocations. The present analysis cannot distinguish among these mechanisms. Similarly, the clearer contemporaneous results for expenditure shares than for defence investment or arms imports might reflect differences in persistence, procurement cycles, implementation delays, reporting practices, or the domestic and foreign composition of purchases. These possibilities were not measured or tested and remain hypotheses rather than identified transmission channels.

The second implication is that conditional association and predictive transfer are distinct empirical properties. The two primary military-expenditure measures had positive coefficients in the fixed-effects models, but none of the complete predictive models demonstrated reliable improvement over the training-mean benchmark. In the matched comparison, Ridge, Elastic Net, and gradient boosting performed better when the defence indicators were omitted, while the Lasso difference remained uncertain. Predictor selection also varied across validation folds. Statistical evidence within a panel-regression specification therefore does not imply that the same variables improve prediction for a country excluded from model fitting. Studies of defence–FDI relationships should report out-of-sample validation separately from explanatory regression results rather than treating model fit, variable retention, or fitted importance as confirmation of an underlying effect.

The temporal analysis adds a related qualification. Conventional Dumitrescu–Hurlin statistics suggested several lagged relationships, but residual dependence remained across countries. Once common cross-country and short-range temporal dependence were incorporated through the moving-block bootstrap, none of the 24 direction–lag combinations remained below the 5% false-discovery-rate threshold. The results therefore

provide no robust basis for interpreting the relationship as delayed, reciprocal, or directionally predictive at the examined annual lags. Future research should continue to allow for reverse directionality and common shocks, but it should not treat the lag orders suggested by the conventional tests as established adjustment periods.

The additional structural and dynamic analyses likewise provide no firm basis for interpreting the contemporaneous associations as stable policy effects. Joint tests did not identify changes in the slopes of the primary indicators after either 2014 or 2022. Lagged-outcome fixed-effects models retained positive coefficients, but the inferential conclusions depended on the covariance estimator, and the difference-GMM specification failed the Hansen test of instrument validity. These results neither demonstrate structural stability nor provide confirmatory dynamic evidence. Stronger causal assessment would require longer time series, credible external sources of variation, and dynamic specifications supported by valid and sufficiently parsimonious instrument sets.

The main policy implication is one of caution. The findings do not justify increasing defence expenditure as an instrument for attracting FDI. The positive panel coefficients represent conditional contemporaneous associations in the main sample; their magnitudes and inferential support varied across countries, periods, scaling choices, and covariance estimators. They were not reproduced in the EU27 extension or supported by transferable predictive performance. Defence policy should therefore be assessed primarily according to security objectives, financing requirements, fiscal constraints, and opportunity costs. From a sustainable-development perspective, the assessment should also consider whether security-related allocations preserve the capacity to finance infrastructure, education, innovation, social provision, and other sources of long-term productive capacity.

The study did not include a direct measure of fiscal sustainability and cannot determine whether a particular level or composition of defence expenditure is fiscally sustainable. Its contribution to this question is limited to showing that defence allocations cannot be evaluated on the assumption that higher expenditure will generally be accompanied by stronger foreign investment. Claims about debt sustainability, fiscal space, expenditure displacement, or long-term budgetary resilience require dedicated fiscal variables and a research design constructed specifically for those outcomes.

The composition of defence expenditure may be more informative for future policy evaluation than its aggregate level. Economic consequences may differ according to whether procurement is domestic or imported, whether it supports research and dual-use technologies, whether local firms participate in supply chains, and whether projects generate civilian technological or infrastructural spillovers. The absence of a general association for defence investment or arms imports does not determine the value of individual programmes. Project-, firm-, and sector-level evidence is required to identify which forms of expenditure strengthen productive capacity, attract productive investment, or primarily create fiscal costs or import demand.

Finally, neither global geopolitical risk nor the three country profiles produced statistically supported differences in the defence–FDI coefficients. However, the absence of detected heterogeneity does not warrant uniform policy prescriptions. The uncertainty measures did not capture differences in national exposure, and the profiles summarised broad long-term country characteristics rather than specific defence-industrial capabilities. Policy assessment should reflect each country's economic structure, fiscal capacity, institutional quality, technological specialisation, security exposure, and domestic production base. It should also distinguish the composition and developmental contribution of foreign investment rather than relying on its aggregate volume, since higher net FDI inflows do not necessarily imply stronger environmental, social, or long-term economic sustainability.

https://doi.org/10.3390/su182110962

4.4. Limitations and Directions for Future Research

The principal limitation remains the absence of causal identification. Country fixed effects account for time-invariant national characteristics, while year fixed effects absorb shocks common to all countries. Nevertheless, the estimates may still reflect omitted time-varying factors affecting both defence activity and FDI. Reverse directionality also cannot be excluded, even though the dependence-robust temporal analysis did not establish adjusted predictive relationships in either direction. The lagged-outcome fixed-effects models did not resolve this problem, and the difference-GMM specifications failed the Hansen tests of instrument validity. Likewise, the absence of statistically supported slope changes after 2014 or 2022 does not establish structural stability, because tests based on two pre-specified break dates may have limited power and cannot capture gradual or country-specific changes. Future research could employ natural experiments or carefully justified instrumental-variable and difference-in-differences designs. Any proposed instrument or security-related policy shocks would require explicit validation because geopolitical events may affect FDI directly.

Cross-sectional dependence remains an important concern. EU economies are connected through trade, financial markets, institutions, and security arrangements, while the three uncertainty indicators assign the same annual value to every country. Driscoll–Kraay inference was used to accommodate general cross-sectional and temporal dependence in the contemporaneous panel regressions, while country-clustered and wild-cluster-bootstrap results provided complementary checks. In the temporal analysis, annual cross-sectional demeaning did not eliminate dependence in the unrestricted-model residuals. The restricted-model residuals were therefore resampled as common annual vectors using a synchronised circular moving-block bootstrap, while the unrestricted-model residuals were used only for the Pesaran CD diagnostic. Cross-sectional demeaning was not repeated within each bootstrap replication because the original series had already been transformed and the synchronised residual vectors were intended to preserve the remaining contemporaneous dependence. This procedure materially weakened the conventional Dumitrescu–Hurlin results, but it cannot represent every possible common-factor, spatial, or network dependence structure. Moreover, the panel contains only 19 years and 23 countries, limiting the precision of all covariance estimators and bootstrap procedures. Future studies could model common factors directly through common correlated effects, interactive fixed-effects, spatial-panel, or related approaches suited to interconnected economies.

Stationarity assessment is also constrained by the short time dimension. The CIPS test accounts for common cross-country factors and provided panel-level evidence against the unit-root null for most variables, but rejection for a panel does not imply that every national series is stationary. Country-specific ADF and KPSS diagnostics showed heterogeneous and sometimes conflicting results, while their power is limited with only 19 annual observations. Log military expenditure per capita was particularly sensitive: its unit-root null was rejected only at the 10% level in the complete sample and was not rejected after Hungary, Ireland, and the Netherlands were excluded. First-differencing addressed this concern in the temporal analysis, but it also changed the interpretation and reduced the available time dimension. Sample-specific evidence also differed for FDI in the smaller balanced panels: the unit-root null was rejected at the 1% level in the 18-country defence-investment sample, but only at the 10% level in the 13-country arms-import sample. The temporal arms-import results were consequently treated as exploratory. Longer series would permit more informative country-specific and second-generation panel diagnostics, structural-break tests, and comparisons of level, difference, and error-correction specifications.

The main sample was restricted to 23 EU countries observed from 2005 to 2023. This improves institutional comparability but limits transferability to countries facing active conflict, economies outside collective-security arrangements, and states with substantial domestic defence industries. The results were also sensitive to country and period coverage. Excluding Hungary, Ireland, and the Netherlands reduced the estimated coefficients of the primary indicators and weakened their inferential support, while the analyses ending in 2021 or excluding 2020–2021 produced positive but inferentially variable results. The available-case EU27 extension did not reproduce the primary findings. Re-extracting the original 23 countries from the EU27 dataset reproduced the principal estimates, indicating that the difference did not arise from altered variable construction. However, complete observations for all 27 countries were available for only 14 common years, and the matched-period estimates remained sensitive to country composition. The EU27 results therefore broaden geographical coverage at the cost of a shorter and less uniform panel. The small number of country clusters also limits cluster-robust inference, particularly in profile-related analyses. Wider samples should preserve explicit institutional, security, and data-availability distinctions rather than treating additional countries as homogeneous observations.

FDI net inflows as a percentage of GDP are highly aggregated. The measure combines new equity investment, reinvested earnings, intracompany debt, restructuring, and disinvestment, while negative and extreme observations may reflect special-purpose entities or pass-through financing. Excluding Hungary, Ireland, and the Netherlands reduced the influence of several unusual country series but could not separate productive investment from financial-routing activity within the remaining observations. The dependent variable also does not identify whether investment supports environmental objectives, social priorities, technological upgrading, or longer-term productive capacity. Scaling FDI and military expenditure by GDP may create a mechanical relationship through a shared or related denominator. The common per-capita specification reduced this concern but produced mixed inferential evidence. Future studies should compare ratio-based and real monetary measures and distinguish greenfield investment, mergers and acquisitions, bilateral flows, investor origin, economic sector, and the developmental characteristics of individual investments.

The defence indicators capture distinct but incomplete dimensions. Aggregate expenditure does not reveal its allocation among personnel, operations, infrastructure, equipment, or research. Defence investment does not distinguish domestic procurement from imports or identify the participation of local and foreign firms. Expressing defence investment relative to GDP or estimated total military expenditure did not provide stronger evidence, but these derived ratios remain aggregate measures and inherit measurement uncertainty from their component series. SIPRI trend-indicator values describe transfers of major conventional weapons but are neither monetary values nor physical counts. More detailed information on expenditure composition, domestic content, defence-industry production, employment, exports, and dual-use research would permit direct examination of industrial and technological mechanisms.

Missing data constrained some analyses. Interpolation was limited to eight internal gaps, while 14 defence-investment observations were retained as missing. Temporal tests for defence investment were consequently based on an 18-country balanced panel. Missing arms-import observations were not interpreted as verified zero transfers, and the corresponding temporal tests used only 13 countries with complete originally observed series. Although this treatment avoids imposing unsupported zero values, it reduces statistical power and may produce a selected sample. Linking SIPRI data with national procurement and transfer records could help distinguish genuinely zero observations from incomplete reporting.

The PCA indices and country profiles are descriptive summaries rather than definitive latent constructs or country types. PCA was applied to pooled country–year observations and did not model dependence between repeated observations from the same country. PC1 captured comparatively limited variance in the economic and human-capital groups, and the human-capital KMO value was below 0.50. Parallel analysis also indicated that more than one component could be retained for the economic and political-institutional groups. The country profiles removed annual variation and may depend on the selected variables, period, scaling, and clustering procedure. Future research could compare theoretically specified indices, multilevel factor models, sparse PCA, longitudinal clustering, and latent-class panel approaches, while reporting uncertainty in both component construction and profile assignment.

The predictive analyses were limited by the small number of countries available for grouped validation. Nested country-grouped cross-validation reduced information leakage by conducting variable screening, preprocessing, PCA estimation, and hyperparameter selection within the training data, but each outer validation fold contained only a small set of countries. The resulting performance estimates therefore remain uncertain and may depend on the particular fold composition. The matched comparisons used identical outer folds and out-of-fold observations for the complete and no-defence predictor sets. Removing the defence indicators improved RMSE for Ridge, Elastic Net, and gradient boosting, whereas the Lasso comparison was inconclusive. However, the country bootstrap resampled countries together with their fixed out-of-fold predictions and did not repeat the complete screening, tuning, and fitting procedure within each bootstrap sample. The reported intervals therefore quantify uncertainty in the observed paired prediction-error differences rather than every source of model-selection uncertainty. The design also evaluated transfer to excluded countries rather than forecasting future years. Larger datasets would permit repeated nested partitions, separate country-based and temporal validation, full-pipeline resampling, and an independent final test sample.

Finally, the temporal analysis tested 24 combinations of defence indicator, direction, and lag order. Four unadjusted bootstrap p-values were below 0.05 under the principal three-year block specification, but none remained below the 5% false-discovery-rate threshold after Benjamini–Hochberg adjustment. The same multiplicity-adjusted conclusion was obtained using block lengths of two and four years. The synchronised moving-block bootstrap addressed residual cross-country and short-range temporal dependence, but its performance is constrained by the short annual series and the limited number of cross-sectional units. The defence-investment and arms-import analyses additionally relied on smaller balanced panels, and the weaker sample-specific stationarity evidence for FDI further limits the arms-import results. Only common one- and two-year lag structures were examined. The tests can indicate predictability for at least a subset of countries but do not identify those countries, coefficient signs, or effect magnitudes. Failure to reject the adjusted bootstrap null should therefore be interpreted as insufficient evidence under the present design rather than proof that no temporal relationship exists. Longer panels would support heterogeneous dynamic models, country-specific diagnostics, richer lag structures, and more detailed examination of adjustment periods.

These limitations do not negate the observed associations, but they restrict their interpretation and generalisability. Future research should move from aggregate country-level relationships towards mechanism-based analysis using disaggregated FDI, procurement, industrial, technological, fiscal, and national security-exposure data. Linking these sources with measures of investment quality, civilian spillovers, public-resource allocation, fiscal sustainability, and environmental or social outcomes would permit a more direct assessment of how security relates to longer-term development. Such research should

combine stronger identification strategies, explicit modelling of common shocks, second-generation time-series diagnostics, and independent predictive validation.

5. Conclusions

This study examined the relationship between defence-sector intensity and FDI in 23 European Union countries from 2005 to 2023. Six defence indicators were assessed using descriptive and dimensionality-reduction methods, nested predictive modelling, fixed-effects panel regressions, interaction tests, temporal-predictability procedures, and additional country, period, measurement, structural, and dynamic sensitivity analyses. These methods addressed different statistical questions and did not produce interchangeable forms of evidence.

The fixed-effects models provided qualified evidence of positive contemporaneous associations between FDI and military expenditure relative to GDP or government expenditure in the main 23-country sample. In the fully adjusted B2 specifications, the tests of both coefficients yielded unadjusted p-values below 0.05 under country-clustered, Driscoll–Kraay, and wild-cluster-bootstrap inference. After Holm adjustment for the two primary hypotheses, support remained under Driscoll–Kraay and wild-cluster-bootstrap inference but not under country-clustered inference. Military expenditure per capita and GMI also produced positive secondary coefficients in the level-based models, whereas defence investment and arms imports did not. The indicator-specific results show that aggregate expenditure, budgetary allocation, per-capita expenditure, investment, militarisation, and arms transfers should not be treated as interchangeable measures of a single defence-sector effect.

The principal associations were sensitive to the analytical sample and variable representation. Excluding Hungary, Ireland, and the Netherlands reduced both estimated magnitudes, and neither coefficient remained below 0.05 under wild-cluster-bootstrap inference. The common per-capita specification produced mixed results, while the first-difference model for military expenditure per capita was unsupported. Neither primary association was reproduced in the available-case EU27 extension. Matched-period analyses showed that this divergence reflected country composition together with the shorter complete observation window rather than differences in variable construction. The panel findings should therefore be interpreted as sample- and specification-dependent contemporaneous associations rather than general or causal effects.

The predictive analyses produced a more restrictive conclusion. None of the models using the complete predictor set demonstrated reliable improvement over the training-mean benchmark for countries excluded from model fitting. In direct paired comparisons, Ridge, Elastic Net, and gradient boosting produced lower out-of-fold errors when all defence indicators were omitted, while the Lasso comparison was inconclusive. Defence-indicator retention also varied across outer training samples. Positive panel coefficients therefore did not translate into stable incremental predictive information or reliable transfer to an unobserved country.

Neither geopolitical risk nor the data-derived country profiles produced statistically supported differences in the defence–FDI coefficients. The conventional Dumitrescu–Hurlin statistics suggested several lagged relationships, but residual dependence remained across countries. Under the restricted-residual circular moving-block bootstrap, four of the 24 unadjusted p-values were below 0.05, but none remained below the 5% false-discovery-rate threshold after Benjamini–Hochberg adjustment. Alternative block lengths produced the same multiplicity-adjusted conclusion. The temporal arms-import results were additionally limited by the 13-country balanced sample and weaker sample-specific stationarity evidence for FDI. The structural-change tests and dynamic specifications likewise did not provide independent confirmation: no joint slope changes were identified

after 2014 or 2022, and the difference-GMM specifications failed the Hansen tests of instrument validity. The combined evidence does not establish temporal direction, structural stability, or causality.

The main conclusion is that defence-sector intensity cannot be assumed either to attract or to displace FDI uniformly. The observed relationship depends on the defence measure, country and period coverage, model specification, inferential procedure, and statistical question. The evidence does not support using higher defence expenditure as an instrument for attracting foreign investment. Defence policy should instead be evaluated according to security objectives, financing constraints, fiscal opportunity costs, and expenditure composition. From a sustainable-development perspective, policy assessment should also consider whether security-related allocations preserve the capacity to invest in infrastructure, education, innovation, social provision, and other sources of long-term productive capacity. However, this study did not measure fiscal sustainability directly and cannot determine whether a particular level of defence expenditure is fiscally sustainable.

The aggregated FDI measure used here does not distinguish productive greenfield investment from acquisitions, corporate restructuring, intracompany financing, pass-through flows, or disinvestment, nor does it identify the environmental, social, technological, or developmental quality of individual investments. Future research should use disaggregated sectoral and bilateral FDI data with detailed information on procurement composition, domestic industrial participation, dual-use research, fiscal conditions, and country-specific security exposure. Longer panels and broader but institutionally differentiated samples would support stronger causal designs, explicit modelling of common shocks, heterogeneous dynamic relationships, second-generation time-series diagnostics, and separate country-based and temporal predictive validation. Such evidence is needed to determine which forms of defence expenditure, if any, contribute to productive investment and economic resilience without undermining other long-term development priorities.

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/su182110962/s1, Table S1A: CIPS panel unit-root test results for the principal sample; Table S1B: CIPS results in the full, reduced-country, and variable-specific temporal samples; Table S1C: Summary of country-level ADF and KPSS diagnostics; Table S1D: First-difference sensitivity for log military expenditure per capita; Table S2A: PCA summary and parallel-analysis results; Table S2B: Loadings on the first principal component; Table S3: Internal validation indices for alternative clustering solutions; Figure S1: Standardised country-profile matrix used in the heterogeneity analysis; Table S4: Pesaran CD tests for the B2 model residuals; Table S5A: Leave-one-country-out estimates for the primary B2 models; Table S5B: Ten observations with the largest Cook's distance in each primary B2 model; Table S5C: Multiplicity-adjusted inference for the two primary B2 coefficients; Table S5D: Driscoll–Kraay bandwidth sensitivity for the two primary B2 coefficients; Table S6A: Overall out-of-fold predictive performance and country-bootstrap confidence intervals; Table S6B: Predictive performance by outer country-grouped fold; Table S6C: Hyperparameters selected within each outer training set; Table S6D: Predictor-selection frequency across the five outer folds; Table S6E: Matched nested out-of-country predictive performance with and without defence variables; Table S6F: Paired country-bootstrap comparisons of out-of-fold predictive losses; Table S7A: Conventional time-demeaned Dumitrescu–Hurlin temporal-predictability results; Table S7B: Dependence diagnostics and synchronous moving-block bootstrap for temporal-predictability tests; Table S8A: Primary B2 estimates under additional country and period sensitivities; Table S8B: Period-specific marginal slopes from the structural-break specifications; Table S8C: Joint tests of the post-2014 and post-2022 slope changes; Table S8D: Marginal defence slopes at selected percentiles of the Geopolitical Risk Index; Table S9A: Descriptive statistics for the alternative

defence-investment ratios; Table S9B: Panel estimates for the alternative defence-investment measures; Table S10: Standardised primary B2 estimates in the main and sensitivity samples; Table S11A: Dynamic fixed-effects sensitivity estimates; Table S11B: Parsimonious one-step difference-GMM sensitivity estimates; Table S12A: Overall out-of-fold predictive performance after excluding three countries; Table S12B: Predictive performance by outer country-grouped fold after excluding three countries; Table S12C: Hyperparameters selected within each outer training sample after excluding three countries; Table S12D: Predictor-selection frequencies after excluding three countries; Table S12E: Fold-specific sparse-model coefficients after excluding Hungary, Ireland, and the Netherlands; Table S13: Complete temporal-predictability results after excluding three countries; Table S14A: Primary B2 estimates across 23-country and EU27 samples; Table S14B: Country-level completeness for the EU27 B2 comparison. Supplementary Data S1 contains the harmonised analytical data, codebook, processing audit, analysis provenance, and reader-facing third-round results. Code S1 contains the Python scripts, tested software requirements, fixed random seeds, execution instructions, and generated analytical outputs required to reproduce the reported statistical and predictive analyses.

Author Contributions: Conceptualization, T.R. and S.R.; methodology, T.R.; software, T.R.; validation, T.R.; formal analysis, T.R. and S.R.; investigation, S.R.; data curation, S.R.; writing—original draft preparation, T.R.; writing—review and editing, T.R. and S.R.; visualization, T.R.; supervision, T.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The harmonised analytical dataset supporting the findings of this study is provided as Supplementary Data S1. The workbook contains the principal 23-country analytical data, the EU27 sensitivity data, derived variables, country-profile information, country-level completeness indicators, data-processing records, and an analysis-provenance index. Its Codebook worksheet reports the definition, unit, analytical role, processing treatment, original source, series or dataset identifier, URL, access date, and reuse information for each variable. The underlying source data remain available from the original providers identified in Section 2.3 and the reference list. Code S1 contains the Python scripts, tested under Python 3.12.14, together with the exact package requirements, fixed random seeds, execution instructions, and generated outputs required to reproduce the reported statistical and predictive analyses.

Acknowledgments: During the preparation of this manuscript, the authors used OpenAI ChatGPT (web application; https://chatgpt.com/; accessed on 26 October 2026) and Google Translate (web application; https://translate.google.com/; accessed on 26 October 2026) solety to assist with English-language editing. The authors reviewed and edited all outputs, verified the analyses and references, and take full responsibility for the content of this publication.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

EU	European Union
FDI	Foreign direct investment
GDP	Gross domestic product
GEPU	Global Economic Policy Uncertainty
GMI	Global Militarisation Index
GPR	Geopolitical Risk

IPS	Im–Pesaran–Shin
KMO	Kaiser–Meyer–Olkin
LLC	Levin–Lin–Chu
PCA	Principal component analysis
PC1	First principal component
RMSE	Root mean squared error

References

1. European Defence Agency. *Defence Data 2023–2024*; European Defence Agency: Brussels, Belgium, 2024. Available online: https://eda.europa.eu/docs/default-source/brochures/1eda---defence-data-23-24---web---v3.pdf (accessed on 5 September 2026).
2. Dunne, J.P.; Tian, N. Military expenditure, economic growth and heterogeneity. *Def. Peace Econ.* **2015**, *26*, 15–31. https://doi.org/10.1080/10242694.2013.848575.
3. Organisation for Economic Co-Operation and Development. *OECD Benchmark Definition of Foreign Direct Investment*, 4th ed.; OECD Publishing: Paris, France, 2008. https://doi.org/10.1787/9789264045743-en.
4. Caldara, D.; Iacoviello, M. Measuring geopolitical risk. *Am. Econ. Rev.* **2022**, *112*, 1194–1225. https://doi.org/10.1257/aer.20191823.
5. Islam, M.S.; Beloucif, A. Determinants of foreign direct investment: A systematic review of the empirical studies. *Foreign Trade Rev.* **2024**, *59*, 309–337. https://doi.org/10.1177/00157325231158846.
6. Kilinc-Ata, N.; Al-Fori, A.M. FinTech Investment, Geopolitical-Economic Uncertainty, and CO_2 Emissions in Low- and Middle-Income Countries: Evidence from Dynamic Panel Models. *J. Risk Financ. Manag.* **2026**, *19*, 362. https://doi.org/10.3390/jrfm19050362.
7. Aziz, N.; Khalid, U. Armed conflict, military expenses and FDI inflow to developing countries. *Def. Peace Econ*. **2019**, *30*, 238–251. https://doi.org/10.1080/10242694.2017.1388066.
8. Douch, M. FDI, defense spending, and economic prosperity. *Bull. Econ. Res.* **2025**, *77*, 86–118. https://doi.org/10.1111/boer.12473.
9. Bhandari, P.; Regmi, N.; Sharma, K. Does conflict reshape the military–FDI nexus? Evidence from a dynamic panel analysis. *Rev. Int. Econ.* **2026**, *34*, 949–966. https://doi.org/10.1111/roie.70062.
10. World Bank. World Bank Open Data. Available online: https://data.worldbank.org/ (accessed on 12 March 2026).
11. Eurostat. Eurostat Database. Available online: https://ec.europa.eu/eurostat/ (accessed on 12 March 2026).
12. Stockholm International Peace Research Institute. SIPRI Military Expenditure Database. Available online: https://doi.org/10.55163/CQGC9685 (accessed on 14 March 2026).
13. European Defence Agency. Defence Data. Available online: https://eda.europa.eu/publications-and-data/defence-data (accessed on 14 March 2026).
14. Bonn International Centre for Conflict Studies. Global Militarisation Index: Categories and Method. Available online: https://gmi.bicc.de/ (accessed on 25 August 2026).
15. Davis, S.J. An index of global economic policy uncertainty. *NBER Working Paper No. 22740; National Bureau of Economic Research: Cambridge, MA, USA.* **2016**. https://doi.org/10.3386/w22740.
16. Ahir, H.; Bloom, N.; Furceri, D. *The World Uncertainty Index*; NBER Working Paper No. 29763; National Bureau of Economic Research: Cambridge, MA, USA, 2022. https://doi.org/10.3386/w29763.
17. Levin, A.; Lin, C.-F.; Chu, C.-S.J. Unit root tests in panel data: Asymptotic and finite-sample properties. *J. Econom.* **2002**, *108*, 1–24. https://doi.org/10.1016/S0304-4076(01)00098-7.
18. Im, K.S.; Pesaran, M.H.; Shin, Y. Testing for unit roots in heterogeneous panels. *J. Econom.* **2003**, *115*, 53–74. https://doi.org/10.1016/S0304-4076(03)00092-7.
19. Pesaran, M.H. A simple panel unit root test in the presence of cross-section dependence. *J. Appl. Econom.* **2007**, *22*, 265–312. https://doi.org/10.1002/jae.951.
20. Jolliffe, I.T.; Cadima, J. Principal component analysis: A review and recent developments. *Philos. Trans. R. Soc. A* **2016**, *374*, 20150202. https://doi.org/10.1098/rsta.2015.0202.
21. Kaiser, H.F. An index of factorial simplicity. *Psychometrika* **1974**, *39*, 31–36. https://doi.org/10.1007/BF02291575.
22. Rousseeuw, P.J. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* **1987**, *20*, 53–65. https://doi.org/10.1016/0377-0427(87)90125-7.
23. Davies, D.L.; Bouldin, D.W. A cluster separation measure. *IEEE Trans. Pattern Anal. Mach. Intell.* **1979**, *PAMI-1*, 224–227. https://doi.org/10.1109/TPAMI.1979.4766909.

24. Calinski, T.; Harabasz, J. A dendrite method for cluster analysis. *Commun. Stat.* **1974**, *3*, 1–27. https://doi.org/10.1080/03610927408827101.
25. Friedman, J.H. Greedy function approximation: A gradient boosting machine. *Ann. Stat.* **2001**, *29*, 1189–1232. https://doi.org/10.1214/aos/1013203451.
26. Hoerl, A.E.; Kennard, R.W. Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics* **1970**, *12*, 55–67. https://doi.org/10.1080/00401706.1970.10488634.
27. Tibshirani, R. Regression shrinkage and selection via the Lasso. *J. R. Stat. Soc. Ser. B* **1996**, *58*, 267–288. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x.
28. Zou, H.; Hastie, T. Regularization and variable selection via the Elastic Net. *J. R. Stat. Soc. Ser. B* **2005**, *67*, 301–320. https://doi.org/10.1111/j.1467-9868.2005.00503.x.
29. Wooldridge, J.M. *Econometric Analysis of Cross Section and Panel Data*, 2nd ed.; MIT Press: Cambridge, MA, USA, 2010.
30. Driscoll, J.C.; Kraay, A.C. Consistent covariance matrix estimation with spatially dependent panel data. *Rev. Econ. Stat.* **1998**, *80*, 549–560. https://doi.org/10.1162/003465398557825.
31. Webb, M.D. Reworking Wild Bootstrap Based Inference for Clustered Errors. Queen's Economics Department Working Paper. 2014, No. 1315. Available online: https://ageconsearch.umn.edu/record/274640/?v=pdf (accessed on 26 October 2026).
32. Pesaran, M.H. General diagnostic tests for cross-sectional dependence in panels. *Empir. Econ.* **2021**, *60*, 13–50. https://doi.org/10.1007/s00181-020-01875-7.
33. Holm, S. A simple sequentially rejective multiple test procedure. *Scand. J. Stat.* **1979**, *6*, 65–70.
34. Benjamini, Y.; Hochberg, Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B* **1995**, *57*, 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x.
35. Arellano, M.; Bond, S. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Rev. Econ. Stud.* **1991**, *58*, 277–297. https://doi.org/10.2307/2297968.
36. Roodman, D. A note on the theme of too many instruments. *Oxf. Bull. Econ. Stat.* **2009**, *71*, 135–158. https://doi.org/10.1111/j.1468-0084.2008.00542.x.
37. Granger, C.W.J. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* **1969**, *37*, 424–438. https://doi.org/10.2307/1912791.
38. Dumitrescu, E.-I.; Hurlin, C. Testing for Granger non-causality in heterogeneous panels. *Econ. Model.* **2012**, *29*, 1450–1460. https://doi.org/10.1016/j.econmod.2012.02.014.
39. Damgaard, J.; Elkjaer, T.; Johannesen, N. What Is Real and What Is Not in the Global FDI Network? *J. Int. Money Finance* **2024**, *140*, 102971. https://doi.org/10.1016/j.jimonfin.2023.102971.
40. Esclanda-Lo, C.G.; Masangkay, G.A.; Ong, C.A.S.; Reyes, R.S. Exploring the evolving determinants of foreign direct investment and the investing country preferences in developing Asian economies using machine learning. In Proceedings of the IFC–ECB–Bank of Spain Conference: External Statistics in a Fragmented and Uncertain World, Madrid, Spain, 12–13 February 2024. Available online: https://www.bis.org/ifc/publ/ifcb62_26.pdf (accessed on 25 August 2026).
41. Darmo, Ľ.; Novák, M.; Kálovec, M.; Rybošová, P. Determinants of the FDI inflow into the Visegrad countries. *Ekon. Čas.* **2020**, *68*, 1057–1080. https://doi.org/10.31577/ekoncas.2020.10.05.
42. Truong, L.D.; Friday, H.S.; Pham, T.D. The effects of geopolitical risk on foreign direct investment in a transition economy: Evidence from Vietnam. *J. Risk Financ. Manag.* **2024**, *17*, 101. https://doi.org/10.3390/jrfm17030101.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.