

# Moving NATO Burden Sharing into the Future: Analyzing Burden Sharing in the Post-Cold War Era

**Leon Runje**

*Ministry of Defense, Republic of Croatia*

**Kristijan Kotarski**

*Faculty of Political Science, Zagreb, Republic of Croatia*

## Abstract

**Background:** This paper examines the determinants of NATO military expenditure in the post–Cold War period from 1990 to 2019. It focuses on whether differences among NATO allies can be better explained by economic size, resource endowments, or broader indicators of national capability. **Objectives:** The study compares five predictors of military expenditure: gross domestic product, gross resources, net resources, the Productive Capacities Index, and an adjusted Composite Index of National Capability that excludes the military component. **Method/Approach:** The analysis uses panel models with country- and year-fixed effects, with standard errors clustered by country. **Results/Findings:** The results show that economic size and resource-endowment indicators are most strongly associated with military expenditure. Gross domestic product, gross resources, and net resources show consistently positive and statistically significant within-country effects. By contrast, the Productive Capacities Index is positive but statistically insignificant, while the adjusted Composite Index of National Capability has limited explanatory power. The findings remain stable across alternative samples and model specifications. **Conclusion:** The evidence suggests that economic scale and resource endowments are more useful for explaining variation in defense outlays among NATO allies than broader composite capability indices.

**Keywords:** NATO; post–Cold War period, burden sharing, military expenditure, joint product model

**JEL classification:** F51, F52

**Paper type:** Research article

**Received:** Mar 19, 2025

**Accepted:** Mar 17, 2026

**Citation:** Runje, L., Kotarski, K. (2026). Moving NATO Burden Sharing into the Future: Analyzing Burden Sharing in the Post-Cold War Era. *Business Systems Research*, 17(1), 154-178.

**DOI:** <https://doi.org/10.2478/bsrj-2026-0008>

## Introduction

It makes a big difference to a country's power position whether it has close allies. Effective alliances act as force multipliers by enhancing military power (e.g., intelligence sharing, reciprocal basing, and joint forces), driving economic benefits (trade, investment, and R&D), and boosting international standing through mutual political commitments. Furthermore, one often neglected benefit of alliances is that they provide a deterrence mechanism that increases the cost of aggression to external powers. Nevertheless, alliances are never an unalloyed good. They also come with downsides, which might at times outweigh their upsides. Among them, we can highlight: the risk exposure when a conflict-prone member drags the rest of the Alliance into a war; the political vulnerability when political commitments are not credibly honored; and the cost of burden sharing, as alliances often include provisions for military contributions (Elefteriu, 2023). The broader relevance of security-related economic pressures is also supported by Pejić Bach et al. (2018), who show that the economic cost of violence is associated with countries' internationalization patterns, suggesting that security conditions can have measurable economic consequences beyond purely military outcomes.

Furthermore, debates over burden sharing can either strengthen or strain alliance cohesion, shaping members' commitment to NATO's overarching security goal. The issue has become even more salient in our era of multiple security threats and an increasingly fragmented global economy. Even though burden sharing has always been a source of occasional friction among NATO allies, the issue has escalated further with the election of Donald Trump as the 45th and now 47th President of the United States. Even though the NATO summit held in Wales in 2014 was finalized by a joint pledge to commit 2% of every member's GDP to defense spending and the goal for allies to allocate at least 20% of their defense budgets towards new equipment purchase and military R&D, burden sharing has remained a point of contention between the North American and European flanks of NATO. Both the Trump and Biden administrations have embraced the issue. All of this came to a head with the recent statement by Donald Trump: "If you don't pay up, I'm not going to protect you", which alluded to the potential watering down of Article 5 provision of the NATO Treaty, which is the collective-defense clause that states that an armed attack against one Ally is considered an attack against all (North Atlantic Treaty Organization, 1949).

In light of these developments, a growing body of scholarship has sought to explain why allies differ in their defense contributions, pointing to a complex interaction of political, strategic, and institutional factors. This paper does not attempt to account for the full range of these determinants. Its contribution is narrower but more precise: to assess whether alternative resource endowment and state capability/capacity indicators can improve the measurement of burden sharing when compared to traditional reliance on economic size indicators alone. The relevance of capacity-based explanations is also consistent with broader systems-oriented research on resilience, which shows that actors differ in their anticipation, coping, and adaptation capabilities when facing external shocks (Ahmić, 2022). Building on Olson's theory of collective action, the analysis compares GDP with gross resources (GR), net resources (NR), the Productive Capacities Index (PCI), and a modified version of the Composite Index of National Capability (CINC no ME) that excludes its military component. In this study, economic capacity is understood as the aggregate ability of a state to finance defense, proxied primarily by the level of gross domestic product (GDP). Using panel models with country and year fixed effects, and controlling for security-related factors such as border exposure, terrorism, non-state conflict, and a proximity-varying threat

term based on Russia's military expenditure combined with the inverse distance from each NATO member's capital to Moscow (an extension of two standard controls already used in the literature, e.g., George & Sandler 2021, 2022, 2024), the study evaluates the extent to which these indicators explain variation in military spending among NATO members.

Building on Olson's theory of collective action and the joint-product model, this study hypothesizes that indicators reflecting economic size and resource endowments, specifically  $\ln$  GDP, GR, and NR, will show the strongest and most consistent relationships with military expenditure. These measures change more regularly and directly reflect a state's ability to fund defense. By contrast, broader capacity/capability indices such as PCI and CINC, excluding military expenditure, are expected to perform less well because they change slowly and capture long-term structural or demographic characteristics that may not influence short-term defense decisions. This hypothesis stems from the idea that economic and resource endowments vary more over time and directly reflect the fiscal space available for defense, whereas PCI and CINC change slowly. As a result, economic scale and resource endowments should provide a more reliable explanation of year-to-year variation in NATO members' military expenditure than composite capacity/capability indices such as PCI and CINC.

The paper is structured as follows: Section 1 presents the main research hypothesis and outlines the paper's structure. Section 2: Theoretical Framework and Literature Review outlines the conceptual foundations of the study and positions the contribution within the existing burden-sharing literature. Section 3: Methodology presents data sources, variable construction, and the two-way fixed-effects research design. Section 4 presents the regression findings for all specifications. Section 5: Discussion interprets the results and relates them to broader debates on NATO burden sharing. Finally, Section 6: Conclusion summarizes the main insights and proposes several avenues for further empirical work.

By focusing on this measurement dimension, the paper complements broader theoretical and policy debates on burden sharing and helps refine the empirical tools available for studying collective defense. In light of these trends, this paper contributes to the academic and policy debate on NATO burden sharing by examining how economic size, resource wealth, and capability/capacity indices predict countries' defense commitments.

## Theoretical Framework

This paper applies the theory of collective action as developed by Mancur Olson and Richard Zeckhauser in their seminal 1966 paper titled "An economic theory of alliances". Furthermore, the paper builds on the joint product model of alliance, a theoretical and conceptual framework developed by a range of authors (Olson & Zeckhauser, 1966; Sandler, 1977; Sandler & Forbes, 1980). The development of this literature strand follows a specific trajectory. It starts with Olson and Zeckhauser (1966) and the pure model of the alliance concept. This concept holds that states contribute to collective efforts in international organizations in exchange for a particular benefit. This benefit is a pure public good, which is, by definition, inexhaustible and non-excludable. In the case of the NATO alliance, this good takes on the form of deterrence under the United States nuclear umbrella. Furthermore, the paper considers the findings of Oneal (1990a; 1990b) and Oneal & Diehl (1994), who, citing Olson and Zeckhauser (1966), describe NATO as a "uniquely privileged group" that provides the pure public good of deterrence. Such an organization is fundamentally characterized by the willingness of one dominant actor, in this case the United States,

to bear the brunt of defense costs, largely because of the non-financial benefits it receives from doing so (Oneal, 1990a, p. 445).

The idea of a uniquely privileged group is based on the assumption, within the theory of collective action, that a country's economic size can be a good predictor of its level of national defense ambition. Namely, the logic behind the assumption is as follows: a larger nominal GDP has been shown to correlate (not perfectly, but well enough) with a larger population, a larger landmass, and, subsequently, greater involvement in international affairs. It is important to stress here that this does not include the measurement of GDP per capita but strictly the nominal size of a country's GDP. A good illustration of this logic is the difference between Luxembourg and Russia. While Luxembourg can claim a much higher GDP per capita, this is not necessarily decisive (though it can be relevant) for its defense ambition, especially because it does not align with its natural size or national power. On the other hand, Russia, with its much larger nominal GDP and lower GDP per capita, is a country with a proportionally higher level of defense ambition. Sandler and Shimizu (2014) raise a valid counterargument here, arguing that GDP per capita matters because a country's greater well-being means it has more to lose from conflict or invasion and will therefore spend more on defense.

In terms of positioning this paper within the framework of the existing literature dedicated to exploring the distribution of burden sharing within NATO, it must be noted that NATO's track record with equitable burden sharing has been documented extensively by authors such as Oneal (1990a), Oneal & Diehl (1994), Hillison (2014), Song (2016), Zyla (2015; 2018) and George and Sandler (2018). Additionally, the strong impact of the United States on defense spending patterns among European states, as demonstrated in the works cited above, is of great importance to this paper. Namely, it indicates that one cannot disregard the significant effect that US contributions to NATO have had on impacting the burden-sharing behavior of its European flank.

The pure model of alliance continued to dominate the literature for several decades until it was called into doubt by a significant innovation. Sandler (1977), as well as Sandler and Forbes (1980), redirected the original model's emphasis on the pure public good (such as deterrence through nuclear weapons) towards the excludable (impure) public good of defense (protection through conventional allied forces located on a specific territory) and other ally-specific benefits, the so-called private goods (national defense specific to each nation state). This was facilitated by the Alliance transitioning in 1967 from its nuclear Mutually Assured Destruction (MAD) doctrine to the Flexible Response Doctrine, in which conventional ground forces became more relevant to Alliance security than nuclear forces. Unlike nuclear forces, conventional troops are spatially constrained and rivalrous. Forces deployed to protect one ally are unavailable to protect another. These impure and private benefits differed fundamentally from the non-rival, non-excludable nature of pure defense. In the case of the public good of defense, they reflected a benefit available to all but not inexhaustible. In contrast, the private good of national defense is often a benefit that is private or nation-specific. Consequently, scholars realized that member states derive a complex mix of incentives from participating in NATO, leading to the development of the joint product model of alliance.

This theoretical shift was a consequence of empirical research conducted from the 1970s to the 1990s, which tested Olson and Zeckhauser's claim that larger states bear disproportionate defense burdens. However, the findings showed that the correlation between GDP and military expenditures weakened after NATO adopted a flexible response in 1967. While early work suggested a positive link (Russett, 1970; Beer 1972; Sandler & Cauley, 1975; Sandler & Forbes, 1980), later studies reported declining or

insignificant correlations (Murdoch & Sandler 1982, 1984; Sandler & Murdoch 1986, 2000; Hansen, Murdoch & Sandler 1990; Oneal & Elrod, 1989; Palmer, 1990; Oneal, 1990b; Khanna & Sandler, 1996). These findings, reinforced by George & Sandler (2018) and summarized in Sandler & Shimizu (2014), highlighted the limits of GDP as a sole predictor of burden sharing. This empirical trajectory gave rise to the joint product model, which reconceptualized alliance contributions as producing multiple outputs: pure public goods like nuclear deterrence, impure public goods such as territorial defense, and private goods tied to ally-specific interests (Van Ypersele de Strihou, 1967; Sandler, 1993; Murdoch & Sandler, 1991; Sandler, 2005). This shift emphasized that burden sharing could no longer be understood purely in terms of GDP, but rather as a mix of non-excludable, excludable, and national benefits derived from defense spending.

In recent years, the burden-sharing debate has been revitalized by major geopolitical shocks such as the election of Donald Trump, the COVID-19 pandemic, and the Russian invasion of Ukraine. These developments spurred a wave of scholarship that has both reaffirmed long-standing patterns of free riding and introduced new methodological approaches.

A first line of new research continues to highlight persistent free-riding dynamics. Kim and Sandler (2020), George and Sandler (2024), and George and Sandler (2022) demonstrate that NATO allies have relied heavily on the defense contributions of larger members, especially those bordering Russia, while smaller and poorer allies have systematically under-contributed. Attempts to correct this via political pressure appear limited; for instance, Becker et al. (2024) find that presidential “shaming” of allies often backfires. Sandler and Shimizu (2014, 45) stress that when defense burdens are compared to NATO allies' national income, (...) After 2002, there is (...) clear evidence that various average benefit shares of the allies no longer match their defense burdens. This means that NATO's cohesiveness has fallen following the war on terror, NATO expansion, and the growing importance of out-of-area missions. They find evidence for the exploitation of the richer members by the poorer ones since 2004 or 2010 (Sandler & Shimizu, 2014, pp. 43 - 45).

A second strand of research has explored alternative ways of measuring contributions beyond the traditional 2% of GDP benchmark. Cooper and Stiles (2021) argue that out-of-area troop deployments, adjusted for population, provide a more accurate indicator of NATO commitment. George & Sandler (2024) apply spatial methods to disaggregate spending into operational categories, showing variation in free-riding across personnel, equipment, and infrastructure. Other scholars push the boundaries of what counts as burden sharing: Balcaen et al. (2023) include intelligence and hybrid-threat capacities; Rubin (2023) proposes a “prioritarian model” linking contributions to wealth and security risk; and McInnis et al. (2024) advocate a multidimensional approach encompassing resilience, energy security, and Ukraine aid. Finally, Blankenship (2023) introduces alliance control theory, emphasizing how U.S. efforts to push allies toward higher spending are conditioned by external threats, allies' military capacity, and America's own resource constraints. This framework underscores the enduring centrality of US power in shaping transatlantic burden-sharing dynamics.

A third line of work situates NATO burden-sharing within the broader European defense context. Studies of EU members' defense demand (George & Sandler, 2021; Becker, 2019, 2021) highlight how fiscal rules, unemployment, and domestic political constraints shape spending, complicating attribution between NATO and EU commitments. Meanwhile, Kimball (2024) shows that NATO Centers of Excellence

function as club goods, creating new channels for allies to contribute outside traditional spending metrics.

Taken together, this literature reflects both continuity and innovation. The persistence of free-riding confirms the relevance of collective action theory, while the push toward new indicators highlights long-term dissatisfaction with narrow GDP-based benchmarks. Against this backdrop, the contribution of this paper is deliberately focused: rather than proposing a new paradigm, it tests whether alternative indicators like gross resources, net resources, the Productive Capacities Index, and a modified CINC provide additional explanatory value for defense spending among NATO allies. In doing so, the study aligns with the trends of methodological innovation while remaining focused on the economic dimension of burden sharing. As the study covers the period from 1990 to 2019, a limitation is that the analysis ends in 2019, thus excluding the dramatic changes following Russia's invasion of Ukraine in 2022 and the renewed burden-sharing debates after Donald Trump's re-election. In addition, some variables (e.g., the Productive Capacities Index, PCI) are available only for part of the period, which reduces the number of observations.

## Methodology

In this section, the paper outlines the research design and methodology, focusing specifically on the predictors of military expenditure (ln ME). This study evaluates five core explanatory variables that represent alternative ways of measuring influence on NATO burden sharing. The main variables under study are: (1) the natural logarithm of gross domestic product (ln GDP), (2) gross resources (GR), (3) net resources (NR), (4) the Productive Capacities Index (PCI), and (5) a CINC variant that excludes military expenditure. The variables GR and NR are operationalized by Beckley (2018b). These five indicators form the central focus of the empirical analysis because each captures a form of influence on military expenditure ranging from economic scale to resource endowments and state capacity/capability. The study aims to compare their relative ability to explain defense burden sharing over time. By clearly distinguishing these key predictors from the control variables used to account for external threat, conflict exposure, and border effects, the methodology aims to provide a transparent evaluation of how different capacity measures perform when estimated in an identical empirical framework.

The paper then interprets the results from the panel models with two-way fixed effects (country and year) and clustered standard errors by country. The paper will test the hypothesis that ln GDP, national Gross resources (GR), and national Net resources (NR) predict the level of national defense ambition/military expenditure ambition of NATO member states. At the same time, the UNCTAD's Productive Capacities Index (PCI) and the Composite Index of National Capability without ME (CINC no ME) are expected to lack statistical significance. Furthermore, the paper tests the following independent variables (border exposure, terrorism, non-state conflict, and a proximity-varying threat term based on Russia's military expenditure combined with the inverse distance from each NATO member's capital to Moscow, for assessing burden sharing).

## Research variables

This section provides the definitions of the independent variables: ln ME, ln GDP, PCI, CINC no ME, GR, as well as NR:

- ln ME – logarithm of Military expenditure - denotes the natural logarithm of military expenditure of an individual NATO member state during a single year.

- In GDP – logarithm of Gross Domestic Product - signifies the natural logarithm of the national GDP (nominal GDP expressed in 2015 USD, the data is available from the World Bank, World Development Indicators dataset) of individual NATO member states.
- GR – Gross resources - is an alternative indicator of defense ambition, calculated as the average of a country's share of NATO-wide GDP and its share of a standardized capacity aggregate. As Beckley (Beckley, 2018b; 2018a, p. 19) notes, gross resources provide a standard way to gauge national power in terms of outputs.
- NR - Net resources - is an alternative indicator of defense ambition that combines a country's relative economic size with its productivity per person, weighting size and efficiency equally; by design, it lowers the importance of sheer population when not matched by commensurate output per capita, yielding a cleaner measure of effective or mobilizable resources (Beckley, 2018b; 2018a, p. 18).
- PCI - the UNCTAD's Productive Capacities Index - this is calculated as a geometric average of eight domains or categories: information and communication technologies (ICTs), structural change, natural capital, human capital, energy, transport, the private sector, and institutions. The categories are selected based on their relevance to conceptual and analytical frameworks for building productive capacities. The PCI is a composite index that summarizes the state of productive capacities in economies worldwide by computing scores ranging from 0 to 100 (UNCTAD Productive Capacities Index Methodological Approach and Results, 2020).
- CINC (no ME) - Composite Index of National Capability without ME - is the composite index which combines military spending (removed to avoid mechanical endogeneity with the dependent variable), troop levels, population, urban population, iron and steel production, and energy consumption (Beckley, 2018a, p. 15; National Material Capabilities (V6.0) – Correlates of War, 2021). This measure is generally computed by summing all observations for each of the 6 capability components in a given year, converting each state's absolute component into a share of the international system, and then averaging across the 6 components. CINC (no ME) is expressed on a 0–1 scale. A value of 0.01 indicates that a country accounted for roughly 1% of total global non-military capabilities in the international system in that year. (National Material Capabilities (V6.0) – Correlates of War, 2021).

The paper also adds the aforementioned control variables:  $\ln$  Border,  $\ln$  TransTerror,  $\ln$  NonstateConflict,  $\ln$  ME\_Russia  $\times$   $\ln$  InvDist\_Russia.

- $\ln$  Border – logarithm of a country's exposed borders - denotes the natural logarithm of a country's exposed borders, where 'exposed borders' refer only to those portions of the national boundary that do not adjoin NATO members and hence are subject to heightened external security risks. NATO allies' exposed borders equal their coastline plus their land boundaries with non-NATO countries, including neutral countries (US Central Intelligence Agency 2019 as referenced in Kim & Sandler, 2020: 405). As an ally's exposed borders increase, a thinning of forces occurs, which, in turn, requires more defense spending. This is an important component of security spending. Data are taken from George & Sandler (2021, 2022) and Kim & Sandler (2024). Because exposed borders are defined relative to NATO membership,  $\ln$  Border varies over time as neighboring countries join the Alliance. Consequently, its coefficient in the TWFE model is

identified from within-country changes in exposure to non-NATO borders rather than from cross-country differences in geography.

- **TransTerror** – Transnational terrorism incidents - this variable counts the number of acts of transnational terror affecting the country each year. The number of transnational terrorist attacks is drawn from the International Terrorism: Attributes of Terrorist Events (ITERATE) dataset, which records observations on transnational terrorist events by venue country during 1968–2019 (George & Sandler, 2022: 9). The transnational terrorism variable is defined as the total number of transnational terrorist incidents experienced at home or abroad by an ally for a given year. A higher value means more exposure to international terrorism. This indicator represents non-state threat pressure and is used to determine whether terrorism intensity influences security. The paper measures NATO allies' terrorism-related benefits using transnational terrorist attacks because such incidents, unlike domestic terrorism, involve multiple countries, pose greater threats to territorial integrity, and carry larger economic consequences. Data are taken from George & Sandler (2021, 2022) and Kim & Sandler (2024).
- **In NonstateConflict** – logarithm of the number of non-state conflicts - logarithm of the measure of non-state conflict intensity according to UCDP (Uppsala Conflict Database Program, 2025). The number of global non-state conflicts is drawn from the Uppsala Conflict Data Program (UCDP) Georeferenced Event Dataset (GED) version 21.1 (Sundberg and Melander 2013 as referenced in Kim & Sandler, 2024). According to the UCDP, non-state conflicts are defined as 'the use of armed force between two organized armed groups, neither of which is the government of a state, which results in at least 25 battle-related deaths in a year' (Högbladh, 2021, 29 as referenced in Kim & Sandler, 2024). Data are taken from George & Sandler (2021, 2022, 2024).
- **ME\_Russia × InvDist\_Russia** - Interaction term measuring threat from Russia. It combines (i) the natural log of Russia's military expenditure (which denotes adversary capability) with (ii) inverse distance to Russia, defined as  $1/(\text{great-circle km from each ally's capital to Moscow})$ . In this specification, a given percentage increase in Russian spending (ln term) has a larger effect the closer an ally is to Moscow (i.e., the higher the inverse distance). Thus, higher values indicate greater "threat proximity" (e.g., Baltic high; Portugal/Canada low). Data are available from George & Sandler (2021, 2022) and Kim & Sandler (2024). Related evidence from European countries also suggests that security-related capacities are unevenly distributed across states, including in the domain of information-security practices (Buljan & Spremić, 2019).

To improve comparability and interpretability, variables with highly skewed distributions (military expenditure, GDP, border exposure, and non-state conflict) are used in logarithmic form. Other indicators (GR, NR, PCI, and CINC no ME) are expressed as shares or indices and kept in their original scale, where log transformation would not provide meaningful interpretation.

Before specifying the empirical equations, it is necessary to clarify the dataset's structure, as this directly affects the number of observations and the panel construction. The panel used in this analysis is unbalanced for two reasons. First, NATO membership was not constant over the 1990–2019 period; countries entered the sample only from the year they joined the Alliance, thereby avoiding inclusion of years when they were not yet subject to NATO defense commitments. Second, data availability differs across variables. Several indicators, most notably PCI and the resource-based measures GR, NR, as well as CINC without the military component,

are not available for all countries or all years. As a result, the number of observations varies across model specifications, which is typical for long, unbalanced panel datasets.

### *Model specification*

The regression analysis is based on a two-way fixed-effects (TWFE) specification with country- and year-fixed effects and standard errors clustered at the country level. Country fixed effects control for all time-invariant characteristics of NATO members, such as geography, long-standing defense preferences, and institutional features, while year fixed effects capture shocks common to all states in a given year, including geopolitical and economic developments. Identification, therefore, comes from within-country changes over time rather than cross-sectional differences. Clustering standard errors by country addresses for serial correlation and heteroskedasticity is standard practice in macro-panel settings. Recent studies on defense demand and NATO burden-sharing commonly employ panel models with country-fixed effects and robust or clustered standard errors (e.g., George & Sandler, 2018; George & Sandler, 2022; Kim & Sandler, 2024). Although the specific estimators vary, ranging from two-way fixed-effects models to spatial and dynamic panel methods, the common rationale is to control for unobserved heterogeneity across member states. The TWFE specification used here follows this core logic and ensures comparability with existing empirical work.

### *Model structure*

We formally specify the empirical strategy as follows. The general TWFE equation is: In

$$ME_{it} = \alpha + \beta_1 X_{it} + \theta' W_{it} + \gamma_i + \delta_t + \varepsilon_{it} \quad (1)$$

where:

- $\ln ME_{it}$  – denotes the natural logarithm of military expenditure for country  $i$  in year  $t$  (dependent variable)
- $\alpha$  – denotes the global intercept (not separately identified in a two-way fixed-effects model)
- $X_{it}$  – denotes the main explanatory variable, rotated across  $\{\ln GDP, GR, NR, PCI, CINC \text{ (no ME)}\}$  depending on the specification
- $\beta_1$  – denotes the coefficient capturing the effect of the main explanatory variable on military expenditure
- $W_{it}$  – denotes the vector of control variables ( $\ln Border, TransTerror, \ln NonstateConflict, ME\_Russia \times InvDist\_Russia$ )
- $\theta'$  – denotes the vector of coefficients corresponding to the control variables
- $\gamma_i$  – signifies the country fixed effects controlling for all time-invariant characteristics of each NATO member
- $\delta_t$  – denotes the year fixed effects capturing shocks and trends common to all member states in year  $t$
- $\varepsilon_{it}$  – denotes the idiosyncratic error term

The control vector is:

$$W_{it} = (\ln Border_{it}, TransTerror_{it}, \ln NonstateConflict_{it}, ME\_Russia_t \times InvDist\_Russia_i) \quad (2)$$

Five main specifications are estimated, with each key variable entered separately as the primary predictor of military expenditure. In this way, the extent to which  $\ln GDP$ ,  $GR$ ,  $NR$ ,  $PCI$ , and  $CINC$ , excluding military expenditure, explain variation in military expenditure is assessed. At the same time, the same set of security controls is maintained across all specifications. The explicit equation (with controls) is

$$\ln ME_{it} = \beta_1 X_{it} + \beta_2 \ln Border_{it} + \beta_3 TransTerror_{it} + \beta_4 ME\_Russia_t \times InvDist\_Russia_i + \beta_5 \ln NonstateConflict_{it} + \gamma_i + \delta_t + \varepsilon_{it}. \quad (3)$$

The regressor  $X_{it}$  is rotated across  $\{\ln GDP_{it}, GR_{it}, NR_{it}, PCI_{it}, CINC (no ME)_{it}\}$  (4)

To avoid multicollinearity among the strongly correlated independent variables (particularly  $\ln GDP$ ,  $GR$ , and  $NR$ ), each is introduced separately as the main predictor in equation (3). This means that the core predictor in equation (3) is replaced sequentially across five alternative independent variables:  $\ln GDP$ ,  $GR$ ,  $NR$ ,  $PCI$ , and  $CINC$  (excluding the military component). Each model, therefore, estimates the association between military expenditure and a single focal indicator at a time, holding constant the same set of controls and using identical country- and year-fixed effects. This rotation allows for a clean, like-for-like comparison of the explanatory power of each capacity measure within the same empirical framework.

Country and year fixed effects,  $\gamma_i$  and  $\delta_t$ , are included, and standard errors are clustered at the country level. Because  $InvDist\_Russia_i$  is time invariant and  $ME\_Russia_t$  varies only over time, their main effects are absorbed by  $\gamma_i$  and  $\delta_t$ , respectively; identification of  $\beta_4$  is provided by their interaction.

Coefficients on logged regressors (e.g.,  $\ln GDP$ ,  $\ln Border$ ,  $\ln NonstateConflict$ ) are elasticities: a 1% change in  $X$  is associated with a  $\beta\%$  change in  $ME$ . Coefficients on regressors in levels (e.g.,  $GR$ ,  $NR$ ,  $PCI$ ,  $CINC (no ME)$ ,  $ME\_Russia \times InvDist\_Russia$ ) are semi-elasticities. For variables expressed as shares, a change of  $\Delta$  alters  $ME$  approximately by

$$100 \times (e^{\beta \cdot \Delta} - 1)\%. \quad (5)$$

On the 0 – 1 scale,  $\Delta = 0.01$  corresponds to a one–percentage–point increase, while on the 0 – 100 scale,  $\Delta = 1$  corresponds to a one–percentage–point increase. For small effects, the approximation  $100 \times \beta \cdot \Delta\%$  applies. Because the model includes both country and year fixed effects, the global intercept is not separately identified and is therefore not reported in the regression tables.

The empirical analysis is based on panel data for NATO member countries (beginning with the year of their NATO membership) over the 1990–2019 period. During this period, NATO underwent several rounds of enlargement, growing from 16 members at the beginning of the 1990s to 29 members by 2019. These successive accession waves, including the 1999, 2004, 2009, and 2017 enlargements, fundamentally shaped the composition of the Alliance and, consequently, the structure of the empirical panel used in this study. Countries enter the dataset only from the year they formally joined NATO, ensuring that observed variation reflects the defense behavior of members subject to Alliance commitments. As a result, the panel's temporal and cross-sectional coverage expands over time, mirroring NATO's evolving membership. The panel structure allows us to track variation both across countries and over time, while controlling for unobserved, time-invariant country-specific characteristics. The estimation uses a two-way fixed effects (TWFE) specification with country fixed effects (to absorb persistent cross-country differences) and year fixed effects (to capture common shocks and trends). Standard errors are clustered at the country level to account for potential within-country serial correlation.

Data for the variables used in the paper are available from the Stockholm International Peace Research Institute – SIPRI Military Expenditure Database (Stockholm Peace Research Institute, 2022b), SIPRI Arms Transfers Database (Stockholm Peace Research Institute, 2022a), the United Nations Conference on

Trade and Development (2020), Correlates of War Project (2021), as well as from the World Bank (2024a, 2024b).

## Results

### Descriptive statistics

Table 1 reports descriptive statistics for the main variables, and Table 2 reports descriptive statistics for the control variables. To assess the stability of the estimates, multicollinearity was examined using the Variance Inflation Factor (VIF). All included variables have VIF values below conventional thresholds, indicating that multicollinearity is not a concern (mean VIF = 2.21). The tables report the arithmetic mean, standard deviation, minimum, maximum, and the number of observations (N). Differences in N arise from data availability (e.g., PCI data are available only for 2000–2016).

Table 1  
Summary statistics for the main variables

| Statistics    | Main variables |         |        |        |         |              |
|---------------|----------------|---------|--------|--------|---------|--------------|
|               | In ME          | In GDP  | GR     | NR     | PCI     | CINC (no ME) |
| <b>Mean</b>   | 8.6626         | 26.5837 | 0.0458 | 0.0462 | 58.8741 | 0.0108       |
| <b>StdDev</b> | 1.9343         | 1.6964  | 0.1052 | 0.1220 | 6.4069  | 0.0199       |
| <b>Min</b>    | 4.2268         | 22.1983 | 0.0006 | 0.0000 | 40.1768 | 0.0002       |
| <b>Max</b>    | 13.5396        | 30.6347 | 0.5262 | 0.6108 | 70.9062 | 0.1049       |
| <b>N</b>      | 650            | 650     | 566    | 650    | 494     | 566          |

Source: Authors' work

The main variables show substantial differences across countries: military expenditures (In ME) and economic size (In GDP) vary by several orders of magnitude, Gross Resources (GR) and Net Resources (NR) are volatile, while PCI points to medium-level productive capacities, and CINC (excluding the military component) highlights the strong concentration of resources in a few states.

Table 2  
Summary statistics for the control variables

| Statistics    | Control variables |             |                     |                           |
|---------------|-------------------|-------------|---------------------|---------------------------|
|               | In Border         | TransTerror | In NonstateConflict | ME_Russia xInvDist_Russia |
| <b>Mean</b>   | 7.7474            | 1.9764      | 6.4643              | 24.6050                   |
| <b>StdDev</b> | 2.4653            | 8.9897      | 0.9133              | 17.9427                   |
| <b>Min</b>    | 0.0000            | 0.0000      | 5.2730              | 0.0000                    |
| <b>Max</b>    | 12.2164           | 181.0000    | 8.0681              | 101.1718                  |
| <b>N</b>      | 650               | 635         | 636                 | 648                       |

Source: Authors' work

The control variables reflect cross-country heterogeneity: border exposure (In Border) ranges from none to very high; transnational terrorism (TransTerror) is generally rare but includes extreme cases; non-state violence (In NonstateConflict) shows moderate differences, while ME\_Russia × InvDist\_Russia indicates wide dispersion in security exposure to Russia.

Following Tables 1 and 2, quartile-based graphics (Q1–Q4) for the key variables, In ME, In GDP, GR, NR, PCI, and CINC no ME (excluding military expenditure) are

presented for the 1990–2019 period. For each year, countries are sorted into four quartile groups: Q1 = bottom 25% of the sample; Q2 = 25th–50th percentile; Q3 = 50th–75th percentile; and Q4 = top 25%. This allows visual tracking of levels (movement across quartiles) and dispersion across countries over time (Figures 1 to 6).

For variables  $\ln$  GDP, PCI, GR, NR, and CINC excluding ME, countries in Q1 can be interpreted as having smaller economies and smaller defense ambitions (in relative, within-sample terms), whereas countries in Q4 are economically larger and have higher defense ambitions. For  $\ln$  ME, the same pattern holds for lower vs. higher military spending (Q1 = lower, Q4 = higher). Because  $\ln$  ME and  $\ln$  GDP are in logs, vertical differences can be read approximately in per cent terms; GR/NR/CINC (excl. mil. comp.) are shares (0–1), and PCI is on a 0–100 scale.

These figures are descriptive and do not imply causality; quartiles are computed within the sample and by year. Inferential conclusions rely on the TWFE regression results.

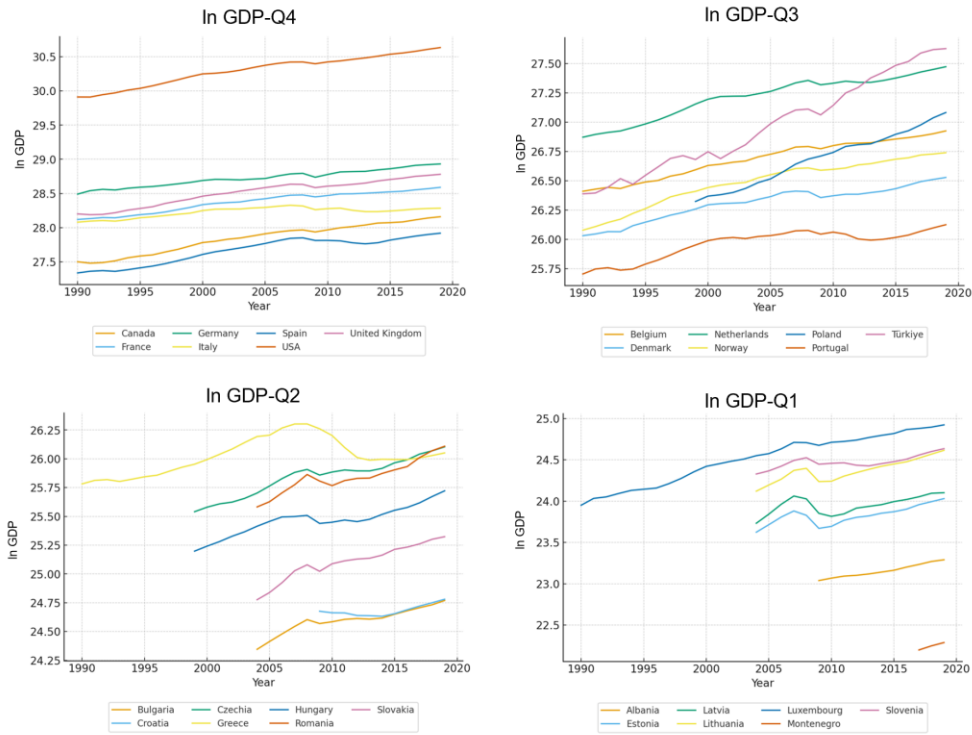
Figure 1

$\ln$  ME – quartile groups by each country's median (static), 1990–2019.



Source: Authors' work

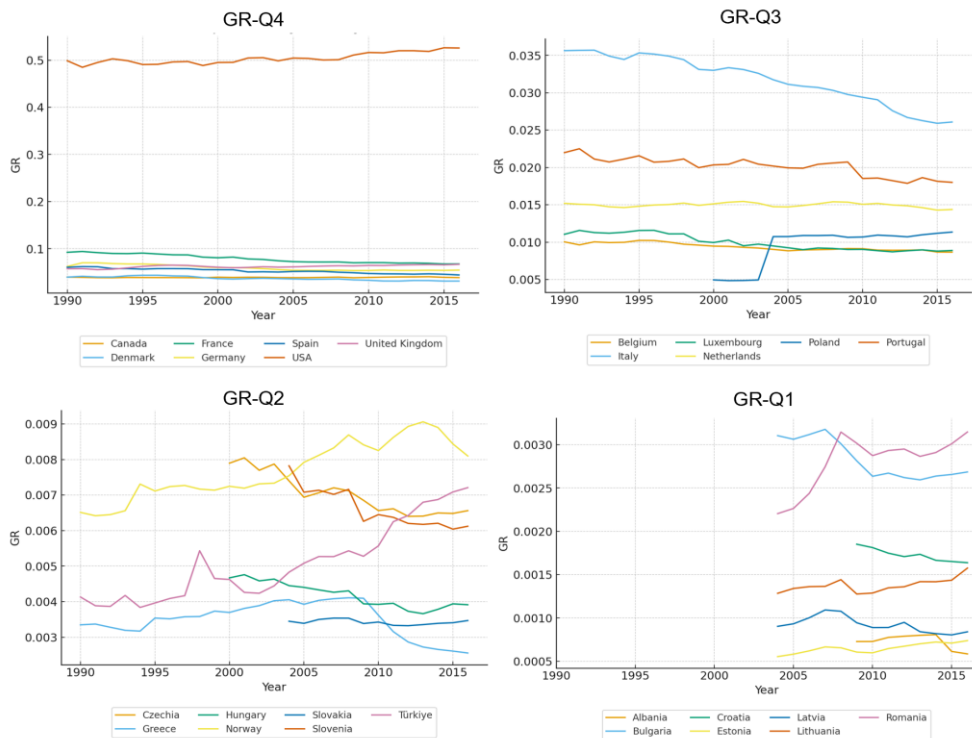
Figure 2  
In GDP- quartile groups by each country's median (static), 1990–2019.



Source: Authors' work

Figure 3

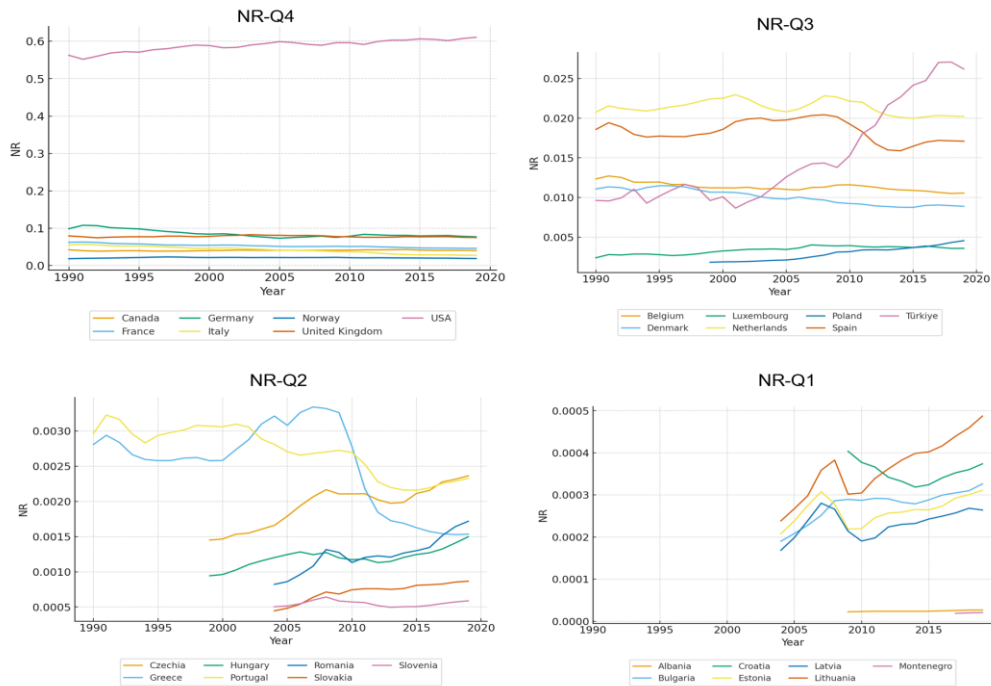
GR- quartile groups by each country's median (static), 1990–2019.



Source: Authors' work

Figure 4

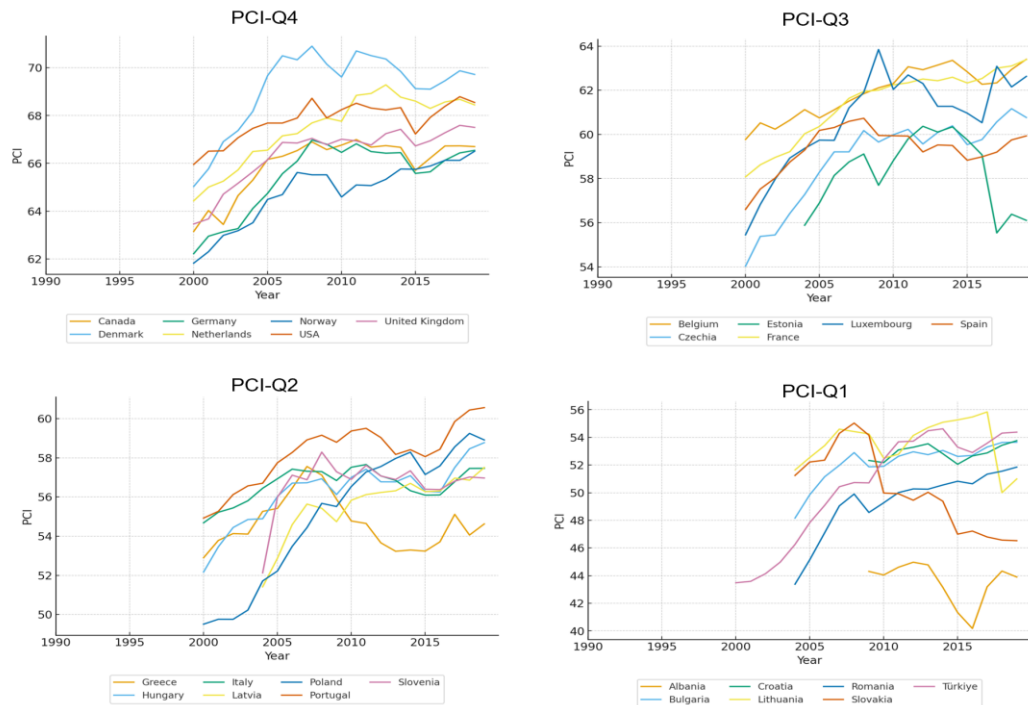
NR– quartile groups by each country's median (static), 1990–2019.



Source: Authors' work

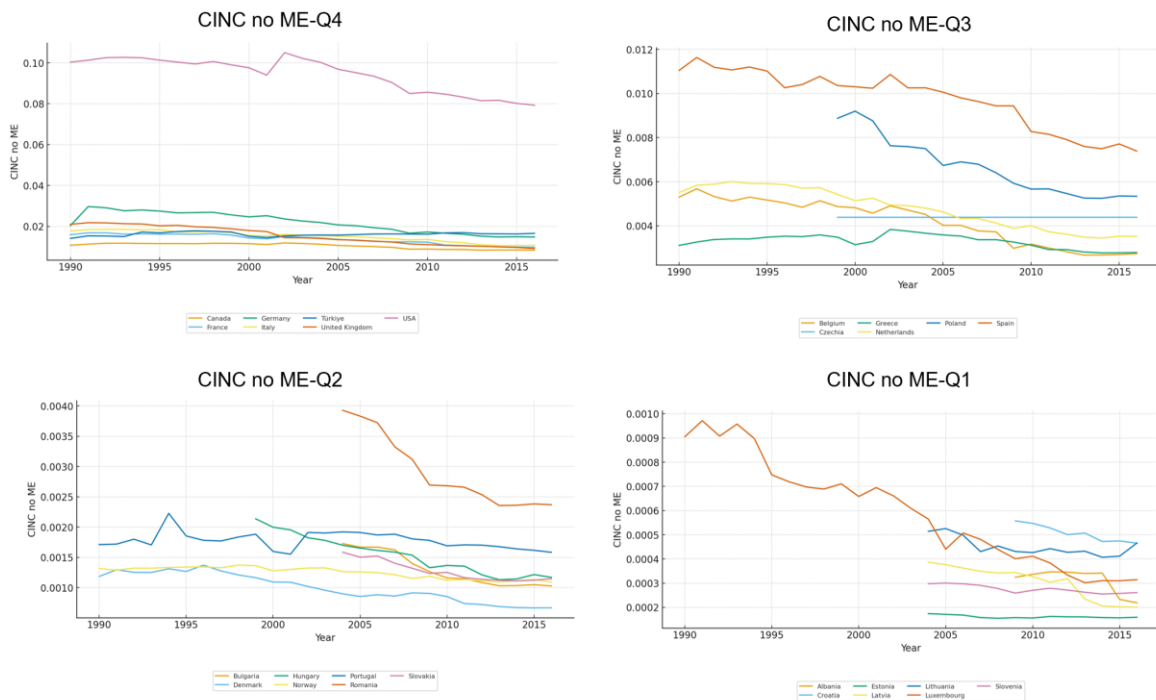
Figure 5

PCI– quartile groups by each country's median (static), 1990–2019.



Source: Authors' work

Figure 6  
CINC (no ME.)– quartile groups by each country's median (static), 1990–2019.



Source: Authors' work

The quartile plots serve as descriptive tools that highlight cross-country heterogeneity and long-run movement in the core variables. Their purpose is not inferential but to illustrate broad trends prior to the regression analysis. The figures show that allies differ substantially in economic scale, resource endowments, and military spending, and that these differences persist over time but also exhibit notable within-country variation, which motivates the subsequent panel analysis.

Taken together, both tables and figures show wide differences across allies and clear movement over time in ME and in key measures (ln GDP, GR, NR, PCI, and CINC no ME). These patterns are informative but remain descriptive: they do not indicate whether year-to-year changes in economic size are associated with year-to-year changes in military spending within countries. To move from descriptive patterns to within-country panel analysis, the analysis proceeds with panel regressions estimated using a two-way fixed-effects (TWFE) framework, as outlined in the Methodology section. This approach exploits within-country, over-time variation while absorbing all country-specific and year-specific shocks. Each of the five indicators is evaluated in an otherwise identical specification that includes the same set of security-related controls (ln Border, TransTerror, ln NonstateConflict, and ME\_Russia × InvDist\_Russia). This harmonized setup enables a clear comparison of which indicators are most strongly associated with changes in military expenditure. The results of the regression analysis are presented in Table 3.

Table 3

Two-Way Fixed Effects (TWFE) model results – dependent variable:  $\ln\_ME$ 

| Variables                                  | Model 1               | Model 2                | Model 3               | Model 4               | Model 5                |
|--|-----------------------|------------------------|-----------------------|-----------------------|------------------------|
| <b>In GDP</b>                              | 0.8369***<br>(0.2784) |                        |                       |                       |                        |
| <b>GR</b>                                  |                       | 12.1926***<br>(2.7809) |                       |                       |                        |
| <b>NR</b>                                  |                       |                        | 9.3997***<br>(1.2685) |                       |                        |
| <b>PCI</b>                                 |                       |                        |                       | 0.0225<br>(0.0145)    |                        |
| <b>CINC (excluding military component)</b> |                       |                        |                       |                       | -10.7750<br>(8.4919)   |
| <b>In Border</b>                           | 0.2517***<br>(0.0584) | 0.2167***<br>(0.0611)  | 0.2349***<br>(0.0586) | 0.2569***<br>(0.0531) | 0.2247***<br>(0.06381) |
| <b>TransTerror</b>                         | 0.0004<br>(0.0004)    | 0.0006<br>(0.0006)     | 0.0001<br>(0.0006)    | 0.0044<br>(0.0036)    | 0.0013**<br>(0.0006)   |
| <b>ME_Russia × InvDist_Russia</b>          | 0.0011<br>(0.0044)    | 0.0024<br>(0.0047)     | 0.0057<br>(0.0050)    | 0.0054<br>(0.0057)    | 0.0023<br>(0.0054)     |
| <b>In NonstateConflict</b>                 | -0.1645<br>(0.1332)   | 0.0060<br>(0.1335)     | 0.0918<br>(0.1034)    | 0.0485<br>(0.0694)    | -0.0507<br>(0.1607)    |
| <b>Year FE</b>                             | Yes                   | Yes                    | Yes                   | Yes                   | Yes                    |
| <b>Country FE</b>                          | Yes                   | Yes                    | Yes                   | Yes                   | Yes                    |
| <b>Observations (N)</b>                    | 634                   | 550                    | 634                   | 494                   | 550                    |
| <b>R-sq within</b>                         | 0.3940                | 0.3030                 | 0.3036                | 0.2903                | 0.2333                 |

Note: Coefficients reported with robust standard errors clustered by country in parentheses. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.10$

Source: Authors' work

By analyzing the impact of various factors on the military expenditure across five regression models, the following results were obtained:

Model 1. Economic size ( $\ln$  GDP) is positively and statistically significantly associated with military expenditure ( $\beta = 0.837$ ,  $SE = 0.278$ ,  $p < 0.01$ ). Interpreted as a within-country elasticity from a two-way fixed-effects model, a 1% increase in GDP is associated with an approximately 0.84% increase in military spending, *ceteris paribus* (95% CI  $\approx 0.27$ – $1.41$ ).

Model 2. GR, expressed as a share (0–1), is positively and statistically significantly associated with military expenditure ( $\beta = 12.193$ ;  $SE = 2.781$ ;  $p < 0.01$ ). In this semi-log specification, a 1 percentage point increase in GR ( $\Delta x = 0.01$ ) is associated with  $\approx 12.9\%$  increase in military spending, *ceteris paribus*. (95% CI  $\approx 6.5$ – $17.9$ ).

Model 3. NR, expressed as a share (0–1), is positively and statistically significantly associated with military expenditure ( $\beta = 9.399$ ;  $SE = 1.269$ ;  $p < 0.01$ ). In this semi-log specification, a 1 percentage point increase in NR ( $\Delta x = 0.01$ ) is associated with  $\approx 9.9\%$  increase in military spending, *ceteris paribus*, (95% CI  $\approx 6.8$ – $12.0$ ). This indicates a sizeable within-country association: countries with greater net resource endowments tend to spend more on the military, consistent with arguments about safeguarding resource rents.

Model 4. PCI is positive but not statistically significant ( $\beta = 0.0225$ ,  $SE = 0.0145$ ,  $p = 0.134$ ). In this semi-log specification ( $\ln$  ME; PCI on a 0–100 scale), a 1-point increase in PCI is associated with an  $\approx 2.3\%$  increase in military spending, *ceteris paribus*. However, the estimate is imprecise; the 95% CI for the 1-point effect is roughly  $-0.6\%$  to  $+5.2\%$ , so the null of a zero effect cannot be rejected. Directionally, the coefficient is consistent

with the idea that stronger productive capacity is associated with higher ME, but the evidence here is not statistically conclusive.

Model 5. CINC (excluding military expenditure) is negative but not statistically significant ( $\beta = -10.775$ ;  $SE = 8.492$ ;  $p = 0.216$ ). Within-country changes in non-military capability are not reliably associated with changes in military spending, *ceteris paribus*. If CINC is on a 0–1 scale, a 1-percentage-point increase ( $\Delta x = 0.01$ ) is associated with an  $\approx 10.2\%$  decrease in ME; the 95% CI is wide ( $\approx -28\%$  to  $+7\%$ ), spanning zero.

In Border. Positive and precisely estimated across all five TWFE models ( $\beta \approx 0.22$ – $0.26$ ;  $SE \approx 0.05$ – $0.07$ ;  $p < 0.01$ ). Interpreted as a within-country elasticity, a 10% increase in the border measure is associated with  $\approx 2.3$ – $2.6\%$  increase in military expenditure, *ceteris paribus*. The magnitude and significance are stable across Models 1–5 (country & year FE; SEs clustered by country).

TransTerror. Coefficients are small and mostly insignificant (one significant at  $p < 0.05$ ). In semi-log terms, a one-unit increase in TransTerror is associated with only a  $\approx 0.01$ – $0.44\%$  increase in ME, suggesting this factor is not a robust driver of military expenditure.

ME\_Russia  $\times$  InvDist\_Russia. Consistently insignificant across all TWFE models, with very small coefficients ( $\beta \approx 0.001$ – $0.006$ ;  $SE \approx 0.004$ – $0.006$ ;  $p \approx 0.26$ – $0.81$ ). In the semi-log setup, a one-unit increase in the interaction variable is associated with only an  $\approx 0.1$ – $0.6\%$  increase in ME; 95% CI spans zero (e.g., Model 3:  $\approx -0.4\%$  to  $+1.6\%$ ). With country and year FEs and clustered SEs, there is no detectable within-country association between ME\_Russia  $\times$  InvDist\_Russia and domestic military spending.

In NonstateConflict. Coefficients are small and imprecise across TWFE models ( $\beta \approx -0.164$  to  $+0.092$ ;  $SE \approx 0.07$ – $0.13$ ;  $p = 0.23$ – $0.92$ ). In this log–log setup, a 10% increase in non-state conflict intensity is associated with roughly a  $-1.6\%$  to  $+0.9\%$  change in ME, and 95% CI spans zero in all models. Conclusion: no reliable within-country association in this sample (country & year FE; SEs clustered by country).

These patterns indicate that the controls are not driving the main results: In Border shows a stable positive association with ME, whereas terrorism, In NonstateConflict, and ME\_Russia  $\times$  InvDist\_Russia do not exhibit robust within-country effects once fixed effects are included. All models include country and year fixed effects, with standard errors clustered by country.

### Robustness checks

To assess robustness, two additional analyses were conducted. First, statistically non-significant control variables were excluded from the specification (Table A1). Second, the models were re-estimated on a restricted sample covering the years 2000–2019 (Table A2). In both cases, the main results remained stable: GDP, GR, and NR were consistently significant, whereas PCI and CINC did not exhibit robust effects. PCI attained significance only at the 10% level when non-significant controls were excluded, and CINC (excluding the military component) attained significance at the 1% level in the restricted sample. These outcomes are interpreted as specification and sample-dependent sensitivities rather than consistent effects. It is therefore confirmed that the main conclusions are not driven by model specification or sample choice. Detailed results are presented in the Appendix.

## Discussion

The results of the regression analysis confirm that economic size and resource-endowment indicators remain central to explaining NATO burden sharing in the post-Cold War era. In GDP, GR, and NR, each displays a robust, statistically significant

positive association with military expenditure. By contrast, PCI is positive but statistically insignificant, while the adjusted CINC index (excluding the military component) lacks explanatory power. This pattern underscores that straightforward measures of economic scale and resource wealth capture much of the variation in defense spending, whereas composite indices of national capacity/capability provide limited additional insight.

These findings resonate with a long line of work that emphasizes the role of economic size in shaping defense burdens (Olson & Zeckhauser, 1966; Oneal, 1990; Khanna & Sandler, 1996). At the same time, they also echo the limitations identified in the literature concerning  $\ln$  GDP as a sole predictor of military effort (Murdoch & Sandler, 1984; Sandler & Murdoch, 2000; George & Sandler, 2018). By incorporating NR and GR, the analysis builds on the argument that net and gross resources can improve our understanding of national power as defined by Beckley (2018b; 2018a) and, as this paper posits, of NATO burden sharing. Our results validate this proposition in the NATO context: both NR and GR perform as strong predictors of defense spending, highlighting that resource efficiency and relative output matter alongside sheer size.

The null results for PCI and CINC (no ME) are equally important. Scholars have increasingly sought to broaden the measurement of burden sharing by moving beyond  $\ln$  GDP. This paper explores multidimensional indicators such as productive capacity or composite capability indices (UNCTAD 2020; Beckley, 2018b; 2018a). The findings suggest that while such indicators may capture broader aspects of state capacity/capability, they do not consistently translate into defense outlays within NATO. This aligns with the view that measures focused on resources or deployments are more closely tied to alliance behavior than abstract indices (Cooper & Stiles, 2021).

Control variables provide further nuance. As this paper finds, border exposure exhibits a consistent positive association with military spending, underscoring that geography and vulnerability remain important drivers of defense effort. This interpretation aligns with the broader literature that highlights the strategic relevance of exposed borders (Sandler, 2005) and spatial proximity to adversaries (George & Sandler, 2022) in Alliance behavior. By contrast, exposure to terrorism, non-state conflict, and Russian military expenditure, when interacting with inverse distance, is not statistically robust.

As shown in the Appendix, the robustness checks confirm that the main findings are stable across alternative specifications and sample restrictions. This further strengthens confidence in the results.

The positive, statistically robust effect of exposed borders reflects changes in border exposure over time, as NATO enlargement altered which borders were classified as exposed. Countries with long coastlines or extensive non-NATO land frontiers must finance persistent patrol, surveillance, and readiness, which persists even after country- and year-level effects are absorbed. By contrast, the interaction between Russia's military expenditure and the inverse distance of a NATO ally's capital to Moscow is insignificant in our two-way fixed effects setting for three plausible reasons. First, most of the variation in Russia's spending is common across countries and is largely captured by year effects, leaving only a country-specific scaling that may be too weak to detect. Second, the interaction between Russia's military expenditure and the inverse distance of a NATO ally's capital to Moscow does not fully capture exposure through Kaliningrad, air and naval reach, or alliance posture. Third, NATO defense budgets adjust slowly and often non-linearly to external shocks, so marginal annual changes in the interaction between Russia's military expenditure and inverse distance of a NATO ally's capital to Moscow need not translate into within-country

spending shifts once fixed effects are included. In line with these mechanisms, the paper finds a clear “level” effect from geography but no reliable “marginal” effect from annual movements in the interaction between Russia's military expenditure and the inverse distance of a NATO ally's capital to Moscow. This suggests that while NATO states recognize these threats, their defense budgets are shaped more by structural resource endowments and territorial exposure than by episodic or transnational risks.

Substantively, the study makes a primary contribution to measurement. It brings five widely used indicators, In GDP, GR, NR, PCI, and CINC, without the military expenditure, into a single, head-to-head framework with identical samples, controls, and estimators and reports effects on comparable elasticity/semi-elasticity scales. The unified estimation, plus transparent diagnostics (stable controls, low VIF, quartile visuals), yields a replicable template that can be extended to other periods, disaggregated defense categories, or additional indicators. The overarching implication is straightforward: in the NATO context, economic and resource-based indicators (economic scale and available resources), as well as geography, better account for defense ambition than broader capability composites.

Taken together, these results make a twofold contribution to the literature. First, they reaffirm the explanatory relevance of economic and resource-based indicators for burden sharing, demonstrating that GR and NR offer valuable refinements and additions to traditional In GDP. Second, they caution against the assumption that more complex indices automatically yield better explanatory power. The limited impact of PCI and CINC (no ME) indicates that policymakers and scholars should treat such measures with caution when assessing allies' contributions.

In terms of originality, the paper makes a distinct, measurement-focused contribution to the burden-sharing literature. First, it brings five leading indicators: In GDP, GR, NR, PCI, and CINC stripped of their military component (CINC no ME) into one harmonized framework and evaluate them under identical conditions: the same sample, the same control set, and a common two-way fixed-effects design with clustered standard errors. This like-for-like setup isolates differences in performance to the indicators themselves rather than to modeling choices. Second, it resolves a frequent comparability problem by removing the military component from CINC to avoid mechanical overlap with the dependent variable and by reporting effects on directly comparable scales, making cross-indicator interpretation straightforward. Third, it adapts GR and NR to the NATO context and tests them against In GDP in a unified panel, showing where resource-based measures add traction over simple economic size. Fourth, it pairs the unified estimators with transparent diagnostics, stable controls, clear interpretations of elasticity/semi-elasticity, and quartile visuals so readers can trace how results map from descriptive patterns to fixed effects. Finally, the study offers a replicable template: the same side-by-side protocol can be extended to new periods, to disaggregated defense categories, or to additional indicators, providing a practical benchmark that researchers and practitioners can reuse when assessing allies' defense contributions.

## Conclusion

This study examined why NATO allies differ in their defense outlays during the post–Cold War period by comparing several indicators of defense ambition within a unified empirical framework with country- and year-fixed effects. The results are consistent and clear. In GDP, gross resources, and net resources are each positively and significantly associated with military expenditure, whereas the Productive Capacities Index and a version of CINC that excludes the military component show little explanatory power. Among the controls, exposed borders show a stable positive

association with military spending, whereas measures of terrorism, non-state conflict, and Russian military expenditure, when interacted with the inverse distance of NATO capitals to Moscow, are not reliably related to within-country changes in defense budgets.

These findings contribute to the burden-sharing literature in two ways. First, they reaffirm the significant role of economic scale and resource endowments in shaping allies' defense efforts, showing that net and gross resources are valid alternatives to In GDP and pointing to a common resource-based logic that holds across specifications. Second, they caution against an assumption that broader composite capability indices will improve the explanation of defense outlays. In the NATO setting considered here, simpler economic and resource measures align more closely with actual spending behavior than multidimensional indices of productive capacity or composite power without military inputs.

The results also carry practical implications. For policy assessment and alliance dialogue, measurement choices matter. Indicators that reflect economic size and resource availability provide a transparent benchmark for evaluating contributions and for diagnosing gaps between expected and observed effort. They help focus debate on economic and resource-based indicators rather than on abstract indices that do not consistently map onto budget decisions.

Future research can extend these insights in several concrete ways. First, it could link size measures, such as GR and NR, to observable outputs, such as deployed troops, equipment procurement, and new or upgraded infrastructure, to see which measure best tracks real defense efforts. Second, instead of measuring burden via military expenditure, define burden using observable outputs such as deployed troops, equipment procurement, and new or upgraded infrastructure, and benchmark GR, NR, and In GDP against PCI and CINC (no ME) on these outcomes to provide an external (outcome-based) validation beyond expenditure models. Third, it could also conduct before-and-after comparisons around 2014 and 2022 to assess how major security-related shocks (the Russian invasion of Ukraine) altered spending patterns. Finally, it could incorporate non-budget contributions more systematically, including training missions, cyber defense, energy resilience, and support to Ukraine. Together, these steps will show more precisely how economic and resource scale translates into tangible alliance contributions.

**Acknowledgement:** Funded by the European Union - NextGenerationEU. The views and opinions expressed are solely those of the author and do not necessarily reflect the official views of the European Union or the European Commission. Neither the European Union nor the European Commission can be held responsible for them.

## References

1. Ahmić, A. (2022). Strategic Sustainability Orientation Influence on Organizational Resilience: Moderating Effect of Firm Size. *Business Systems Research Journal*, 13(1), 169-191. <https://doi.org/10.2478/bsrj-2022-0011>
2. Balcaen, P., Du Bois, C., & Buts, C. (2023). Sharing the Burden of Hybrid Threats: Lessons from the Economics of Alliances. *Defence and Peace Economics*, 34(2), 142-159. <https://doi.org/10.1080/10242694.2021.1991128>
3. Becker, J. M. (2019). Accidental rivals? EU fiscal rules, NATO, and transatlantic burden-sharing. *Journal of Peace Research*, 56(5), 697-713. <https://doi.org/10.1177/0022343319829690>
4. Becker, J. (2021). Rusty guns and buttery soldiers: unemployment and the domestic origins of defense spending. *European Political Science Review*, 13(3), 307-330. <https://doi.org/10.1017/s1755773921000102>

5. Becker, J., Kreps, S. E., Poast, P., & Terman, R. (2024). Transatlantic Shakedown: Presidential Shaming and NATO Burden Sharing. *Journal of Conflict Resolution*, 68(2-3), 195-229. <https://doi.org/10.1177/00220027231167840>
6. Beckley, M. (2018a). The Power of Nations: Measuring What Matters. *International Security*, 43(2), 7-44. [https://doi.org/10.1162/isec\\_a\\_00328](https://doi.org/10.1162/isec_a_00328)
7. Beckley, M. C. (2018b). *Unrivaled: Why America Will Remain the World's Sole Superpower*. Ithaca and London: Cornell University Press.
8. Beer, F. A. (1972). *The Political Economy of Alliances: Benefits, Costs, and Institutions in NATO*. Beverly Hills, CA: Sage Publications.
9. Blankenship, B. D. (2023). *The Burden-Sharing Dilemma*. Cornell University Press.
10. Buljan, A., & Spremić, M. (2019). Cluster Analysis of IT Security Risks in Chosen Sectors. *ENTRENOVA-ENTERPRISE RESEARCH INNOVATION JOURNAL*, 5(1), 489-496.
11. Cooper, S., & Stiles, K. W. (2021). Who commits the most to NATO? It depends on how we measure commitment. *Journal of Peace Research*, 58(6), 1194-1206. <https://doi.org/10.1177/0022343320980820>
12. Correlates of War project. (2021). Correlates of War Data Sets. Retrieved from <https://correlatesofwar.org/data-sets/>
13. Elefteriu, G. (2023). Why Alliances Matter. Retrieved from <https://www.geostrategy.org.uk/app/uploads/2023/12/GSPE02.Why-alliances-matter.pdf>
14. George, J., & Sandler, T. (2024). A spatial analysis of NATO burden sharing at the operational levels. *Kyklos*, 77(4), 1026-1047. <https://doi.org/10.1111/kykl.12401>
15. George, J., & Sandler, T. (2018). Demand for military spending in NATO, 1968-2015: A spatial panel approach. *European Journal of Political Economy*, 53, 222-236. <https://doi.org/10.1016/j.ejpoleco.2017.09.002>
16. George, J., & Sandler, T. (2021). EU Demand for Defense, 1990-2019: A Strategic Spatial Approach. *Games*, 12(1), 13. <https://doi.org/10.3390/g12010013>
17. George, J., & Sandler, T. (2022). NATO defense demand, free riding, and the Russo-Ukrainian war in 2022. *Journal of Industrial and Business Economics*, 49(4), 783-806. <https://doi.org/10.1007/s40812-022-00228-y>
18. Hansen, L., Murdoch, J. C., & Sandler, T. (1990). On distinguishing the behavior of nuclear and non-nuclear allies in NATO. *Defence Economics*, 1(1), 37-55. <https://doi.org/10.1080/10430719008404649>
19. Hillison, J. (2014). Stepping Up: Burden Sharing by NATO's Newest Members. Retrieved from <https://press.armywarcollege.edu/monographs/943>
20. Khanna, J., & Sandler, T. (1996). Nato burden sharing: 1960-1992. *Defence and Peace Economics*, 7(2), 115-133. <https://doi.org/10.1080/10430719608404846>
21. Kim, W., & Sandler, T. (2020). NATO at 70: Pledges, Free Riding, and Benefit-Burden Concordance. *Defence and Peace Economics*, 31(4), 400-413. <https://doi.org/10.1080/10242694.2019.1640937>
22. Kim, W & Sandler, T., (2024). NATO Security Burden Sharing, 1991–2020 *Defence and Peace Economics*, 35(3), 265–280, <https://doi.org/10.1080/10242694.2023.2230408>
23. Kimball, A. L. (2024). Reconceptualising burdens - NATO centres of excellence: club goods, informal institutions, and partner contributions. *Defence Studies*, 24(4), 559-600. <https://doi.org/10.1080/14702436.2024.2380886>
24. McInnis, K., Fata, D., Jensen, B., & Macias, J. M., III. (2024, February 22). *Pulling their weight: The data on NATO responsibility sharing*. Center for Strategic and International Studies. <https://www.csis.org/analysis/pulling-their-weight-data-nato-responsibility-sharing>.
25. Murdoch, J. C., & Sandler, T. (1982). A Theoretical and Empirical Analysis of NATO. *Journal of Conflict Resolution*, 26(2), 237-263. <https://doi.org/10.1177/0022002782026002003>
26. Murdoch, J. C., & Sandler, T. (1984). Complementarity, free riding, and the military expenditures of NATO allies. *Journal of Public Economics*, 25(1-2), 83-101. [https://doi.org/10.1016/0047-2727\(84\)90045-8](https://doi.org/10.1016/0047-2727(84)90045-8)

27. Murdoch, J. C., & Sandler, T. (1991). NATO Burden Sharing and the Forces of Change: Further Observations. *International Studies Quarterly*, 35(1), 109. <https://doi.org/10.2307/2600391>
28. North Atlantic Treaty Organization. (1949). *The North Atlantic Treaty*. [https://www.nato.int/cps/en/natohq/official\\_texts\\_17120.htm](https://www.nato.int/cps/en/natohq/official_texts_17120.htm)
29. Olson, M., & Zeckhauser, R. (1966). An Economic Theory of Alliances. *The Review of Economics and Statistics*, 48(3), 266. <https://doi.org/10.2307/1927082>
30. ONeal, J. R. (1990a). Testing the Theory of Collective Action: Nato Defense Burdens, 1950-1984. *Journal of Conflict Resolution*, 34(3), 426-448. <https://doi.org/10.1177/0022002790034003003>
31. ONeal, J. R. (1990b). The theory of collective action and burden sharing in NATO. *International Organization*, 44(3), 379-402. <https://doi.org/10.1017/s0020818300035335>
32. ONeal, J. R., & Diehl, P. F. (1994). The Theory of Collective Action and NATO Defense Burdens: New Empirical Tests. *Political Research Quarterly*, 47(2), 373. <https://doi.org/10.2307/449016>
33. ONeal, J. R., & Elrod, M. A. (1989). NATO Burden Sharing and the Forces of Change. *International Studies Quarterly*, 33(4), 435. <https://doi.org/10.2307/2600521>
34. Palmer, G. (1990). Corraling the Free Rider: Deterrence and the Western Alliance. *International Studies Quarterly*, 34(2), 147. <https://doi.org/10.2307/2600706>
35. Pejić Bach, M., Dumičić, K., Jaković, B., Nikolić, H., & Žmuk, B. (2018). Exploring impact of economic cost of violence on internationalization: Cluster analysis approach. *International Journal of Engineering Business Management*, 10. <https://doi.org/10.1177/1847979018771244>
36. Rubin, D. (2023). Transcending Two Percent: Toward a Prioritarian Model of NATO Burden-Sharing. *Central European Journal of International and Security Studies*, 17(4), 38–62. <https://doi.org/10.2307/1927082>
37. Russett, B. M. (1970). *What Price Vigilance?: The Burdens of National Defense*. New Haven and London: Yale University Press.
38. Sandler, T. (1977). IMPURITY OF DEFENSE: AN APPLICATION TO THE ECONOMICS OF ALLIANCES. *Kyklos*, 30(3), 443-460. <https://doi.org/10.1111/j.1467-6435.1977.tb02203.x>
39. Sandler, T. (1993). The Economic Theory of Alliances: A Survey. *Journal of Conflict Resolution*, 37(3), 446-483. <https://doi.org/10.1177/0022002793037003003>
40. Sandler, T. (2005). NATO BENEFITS, BURDENS AND BORDERS: COMMENT. *Defence and Peace Economics*, 16(4), 317-321. <https://doi.org/10.1080/10242690500083709>
41. Sandler, T., & Cauley, J. (1975). On the Economic Theory of Alliances. *Journal of Conflict Resolution*, 19(2), 330-348. <https://doi.org/10.1177/002200277501900207>
42. Sandler, T., & Forbes, J. F. (1980). BURDEN SHARING, STRATEGY, AND THE DESIGN OF NATO. *Economic Inquiry*, 18(3), 425-444. <https://doi.org/10.1111/j.1465-7295.1980.tb00588.x>
43. Sandler, T., & Murdoch, J. C. (1986). Defense Burdens and Prospects for the Northern European Allies. In D. B. H. Denoon (Ed.), *Constraints on Strategy: The Economics of Western Security* (pp. 59–113). Washington, DC: Pergamon-Brassey.
44. Sandler, T., & Murdoch, J. C. (2000). On Sharing NATO Defence Burdens in the 1990s and Beyond. *Fiscal Studies*, 21(3), 297-327. <https://doi.org/10.1111/j.1475-5890.2000.tb00026>.
45. Sandler, T., & Shimizu, H. (2014). NATO Burden Sharing 1999-2010: An Altered Alliance. *Foreign Policy Analysis*, 10(1), 43-60. <https://doi.org/10.1111/j.1743-8594.2012.00192.x>
46. Song, Y. (2016). *The US Commitment to NATO in the Post-Cold War Period*. Cham: Springer International Publishing. <https://doi.org/10.1007/978-3-319-33548-3>
47. Stockholm Peace Research Institute (2022a). SIPRI Arms Transfers Database. Retrieved from <https://www.sipri.org/databases/armstransfers>
48. Stockholm Peace Research Institute (2022b). SIPRI Military Expenditure Database. Retrieved from <https://www.sipri.org/databases/milex>
49. The North Atlantic Treaty. (1949, April 4). NATO.

50. United Nations Conference on Trade and Development. (2020). Productive Capacities Index: Methodological Approach and Results. Retrieved from [https://unctad.org/system/files/official-document/aldc2020d3\\_en.pdf](https://unctad.org/system/files/official-document/aldc2020d3_en.pdf)
51. Uppsala Conflict Data Program (2025). Uppsala Conflict Data Program Dataset Download Center. Retrieved from <https://ucdp.uu.se/downloads>
52. van Ypersele de Strihou, J. (1967). Sharing the Defense Burden Among Western Allies. *The Review of Economics and Statistics*, 49(4), 527. <https://doi.org/10.2307/1928338>
53. World Bank. (2024a). World Development Indicators -- GDP (constant 2015 USD). Retrieved from <https://databank.worldbank.org/reports.aspx?source=2&series=NY.GDP.MKTP.KD&country=>
54. World Bank. (2024b). World Development Indicators - GDP per capita (constant 2015 USD). Retrieved from <https://databank.worldbank.org/reports.aspx?source=2&series=NY.GDP.PCAP.KD&country=>
55. Zyla, B. (2015). *Sharing the burden?: NATO and its second-tier powers*. Toronto. Buffalo and London: University of Toronto Press.
56. Zyla, B. (2018). Transatlantic burden sharing: suggesting a new research agenda. *European Security*, 27(4), 515-535. <https://doi.org/10.1080/09662839.2018.1552142>

## Appendix

This Appendix reports the robustness checks. Table A1 presents the TWFE models excluding non-significant control variables, while Table A2 re-estimates the models on a restricted sample (2000–2019). The results confirm that the main findings for In GDP, GR, and NR remain stable across these alternative specifications, while PCI and CINC no ME display only specification- and sample-dependent sensitivity.

Table A1

TWFE model results – robustness check (excluding non-significant controls)

| Variables        | Model 1               | Model 2                | Model 3               | Model 4               | Model 5               |
|------------------|-----------------------|------------------------|-----------------------|-----------------------|-----------------------|
| In GDP           | 0.8632***<br>(0.2704) |                        |                       |                       |                       |
| GR               |                       | 11.0002***<br>(2.0818) |                       |                       |                       |
| NR               |                       |                        | 7.9722***<br>(1.6854) |                       |                       |
| PCI              |                       |                        |                       | 0.0275*<br>(0.0155)   |                       |
| CINC (no ME)     |                       |                        |                       |                       | -3.0738<br>(7.8301)   |
| In Border        | 0.2524***<br>(0.0602) | 0.2283***<br>(0.0681)  | 0.2126***<br>(0.0691) | 0.2369***<br>(0.0545) | 0.2307***<br>(0.0768) |
| Year FE          | Yes                   | Yes                    | Yes                   | Yes                   | Yes                   |
| Country FE       | Yes                   | Yes                    | Yes                   | Yes                   | Yes                   |
| Observations (N) | 650                   | 563                    | 650                   | 494                   | 650                   |
| R-sq within      | 0.3944                | 0.2854                 | 0.2802                | 0.2763                | 0.2384                |

Note: All non-significant control variables excluded; only In Border retained as consistently significant; Source: Authors' work

Table A2

TWFE model results – robustness check (restricted sample, 2000–2019)

| Variables                     | Model 1               | Model 2                | Model 3                | Model 4               | Model 5                |
|-------------------------------|-----------------------|------------------------|------------------------|-----------------------|------------------------|
| In GDP                        | 0.9327***<br>(0.3251) |                        |                        |                       |                        |
| GR                            |                       | 15.3502***<br>(3.8434) |                        |                       |                        |
| NR                            |                       |                        | 12.7481***<br>(2.9614) |                       |                        |
| PCI                           |                       |                        |                        | 0.0225<br>(0.0145)    |                        |
| CINC (no ME)                  |                       |                        |                        |                       | -14.7197**<br>(6.4701) |
| In Border                     | 0.2661***<br>(0.0566) | 0.2253***<br>(0.0698)  | 0.2569***<br>(0.0570)  | 0.2569***<br>(0.0531) | 0.2292***<br>(0.0593)  |
| TransTerror                   | 0.0012<br>(0.0020)    | 0.0012<br>(0.0027)     | 0.0017<br>(0.0027)     | 0.0044<br>(0.0036)    | 0.0029<br>(0.0043)     |
| ME_Russia ×<br>InvDist_Russia | 0.0026<br>(0.0047)    | 0.0039<br>(0.0054)     | 0.0077<br>(0.0053)     | 0.0054<br>(0.0057)    | 0.0096<br>(0.0059)     |
| In NonstateConflict           | -0.0365<br>(0.0563)   | -0.0171<br>(0.0793)    | 0.0768<br>(0.0574)     | 0.0485<br>(0.0694)    | 0.0343<br>(0.0719)     |
| Year FE                       | Yes                   | Yes                    | Yes                    | Yes                   | Yes                    |
| Country FE                    | Yes                   | Yes                    | Yes                    | Yes                   | Yes                    |
| Observations (N)              | 497                   | 413                    | 497                    | 494                   | 497                    |
| R-sq within                   | 0.4006                | 0.2817                 | 0.3037                 | 0.2903                | 0.2875                 |

Note: Models re-estimated on the restricted sample covering 2000–2019; Source: Authors' work

## About the authors

Leon Runje holds a master's degree in political science from the Faculty of Political Science at the University of Zagreb. He has been employed at the Ministry of Defense of the Republic of Croatia since 2019. Currently, he serves as the Head of the Defense Analysis Unit within the Defense Policy Directorate of the Ministry of Defense. He is enrolled in the doctoral program in Political Science at the Faculty of Political Science, University of Zagreb, specializing in Public Policy. The focus of his doctoral research is burden sharing within NATO and the EU. The author can be contacted at: [leonrunje91@gmail.com](mailto:leonrunje91@gmail.com).

Kristijan Kotarski is an Associate Professor in International Political Economy and Economic Policy. His research focuses on the political economy of EU integration, global finance and geopolitics. He is Head of the Centre for European Studies at the Faculty of Political Science, University of Zagreb. His research appeared in journals such as: Europe-Asia Studies, Journal of Contemporary European Studies, China Economic Journal, Review of Radical Political Economics, etc. His most recent book is Corona Economics: The Five Riders of the Apocalypse (2021), coauthored by Velimir Šonje and published by European Liberal Forum, Brussels. The author can be contacted at: [kristijan.kotarski@fpzg.hr](mailto:kristijan.kotarski@fpzg.hr).