sustainability

MDPI

Article

From Commitments to Outcomes: How the Globalisation Implementation Gap Shapes SDG Trade-Offs and the Role of Governance

Oksana Liashenko [1,2,*], Oksana Adamyk [2], Łukasz Skowron [3], Grzegorz Hajduk [4], Oleksandr Dluhopolskyi [5,6,*], Olena Mykhailovska [7] and Nataliia Husarevych [8]

1 Faculty of Economics and Management, Lesya Ukrainka Volyn National University, 43-025 Lutsk, Ukraine
2 Loughborough Business School, Loughborough University, Loughborough LE11 3TU, UK; o.o.adamyk@lboro.ac.uk
3 Faculty of Management, Lublin University of Technology, 20-618 Lublin, Poland; l.skowron@pollub.pl
4 Faculty of Economics and Finance, University of Rzeszów, 35-601 Rzeszow, Poland; ghajduk@ur.edu.pl
5 Institute of Public Administration and Business, WSEI University, 20-209 Lublin, Poland
6 Faculty of Economics and Management, West Ukrainian National University, 46-027 Ternopil, Ukraine
7 Faculty of Management, Finance and Marketing, State University "Kyiv Aviation Institute", 03-058 Kyiv, Ukraine; m-olena@ukr.net
8 Office of International and Project Activities, Kyiv National University of Technologies and Design, 01-011 Kyiv, Ukraine; asshattan@ukr.net
* Correspondence: o.liashenko@lboro.ac.uk (O.L.); o.dluhopolskyi@wunu.edu.ua (O.D.)

Abstract

This study revisits the globalisation–sustainability nexus by focusing on the divergence between formal policy commitments and realised integration. While the 2030 Agenda assumes coherence across the seventeen Sustainable Development Goals (SDGs), empirical evidence increasingly points to uneven progress and systemic trade-offs. We introduce the concept of the globalisation implementation gap, defined as the normalised difference between de jure commitments and de facto integration, and examine its impact on SDG performance. Using a panel dataset of 101 countries over the period of 2000–2023 (N = 2197), we construct gap indices based on the KOF Globalisation framework and estimate two-way fixed-effects models with Driscoll–Kraay standard errors. The results reveal a significant negative association between implementation gaps and overall SDG performance. However, this relationship is highly heterogeneous across goals. Larger gaps are strongly negatively associated with social outcomes, particularly poverty reduction (SDG 1), clean water and sanitation (SDG 6), and sustainable cities (SDG 11), while being positively associated with certain environmental outcomes, including responsible consumption (SDG 12) and apparent biodiversity protection (SDG 15)—the latter, however, is interpreted as associational only, given evidence of reverse causation in placebo lead-testing. Further analysis using panel threshold regression demonstrates that governance quality moderates this relationship non-linearly. Above a critical threshold of institutional quality, the magnitude of the negative gap–SDG association increases substantially, suggesting that stronger governance amplifies rather than buffers the association. In addition, Gaussian mixture clustering identifies three distinct country archetypes (Balanced, Informal, and Policy-Led Integrators). A diagnostic placebo lead test indicates that the SDG 15 association is associational rather than causal (likely reverse causation), and the result for SDG 15 is therefore reported with that caveat. Each is characterised by different gap structures and sustainability trade-offs. Overall, the findings shift the perspective from a uniform globalisation–SDG relationship to a goal-specific and governance-contingent framework. The study highlights the importance of aligning policy commitments with actual integration processes and provides policy-relevant insights for managing SDG trade-offs under

Check for updates

Academic Editor: Richard Ross Shaker

Received: 14 April 2026
Revised: 2 May 2026
Accepted: 4 May 2026
Published: 12 May 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.

conditions of globalisation. The headline gap–SDG association is conditional on the listwise sample of 101 countries with complete data and on Driscoll–Kraay variance estimation; sensitivity analyses show that the result attenuates under multiple imputation on the full 208-country panel and loses statistical significance under cluster-robust standard errors, while the political-channel decomposition and the governance threshold remain robust under both stress tests.

Keywords: SDG trade-offs; goal-level heterogeneity; globalisation implementation gap; de facto vs. de jure globalisation; governance moderation; panel threshold regression; policy coherence; sustainable development; KOF Globalisation Index; country typology

1. Introduction

The 2030 Agenda for Sustainable Development, adopted by the United Nations General Assembly in September 2015, represented an ambitious collective pledge to eradicate poverty, protect the planet, and ensure prosperity for all [1]. A decade on, progress towards the seventeen Sustainable Development Goals (SDGs) has been slower than anticipated, with the Sustainable Development Report 2024 documenting stagnation in global SDG performance since 2020 and mounting evidence that most targets will not be met by the 2030 deadline [2]. Among the structural forces most frequently invoked in both policy discourse and academic research as determinants of SDG attainment is globalisation, broadly understood as the intensification of cross-border flows of goods, capital, people, and ideas, alongside the institutional arrangements that govern them [3,4].

A substantial body of empirical literature has examined the relationship between globalisation and various dimensions of sustainable development, including economic growth [5], environmental degradation [6,7], income inequality [8], and composite SDG performance [9]. Yet the findings remain mixed and often contradictory. One plausible explanation for this ambiguity lies in the multidimensional nature of globalisation itself. Following the seminal revision by S. Gygli et al. [10], the KOF Globalisation Index now distinguishes between de facto globalisation—measuring actual cross-border flows and activities, and de jure globalisation—capturing the policies, treaties, and institutional frameworks designed to facilitate such flows. This distinction matters enormously. A country may sign numerous bilateral investment treaties and trade agreements (high de jure score) whilst experiencing limited actual integration into global markets (low de facto score), or vice versa. D. Quinn et al. [11] demonstrated that the choice between de facto and de jure indicators of financial openness yields systematically different findings in growth regressions, and subsequent work has confirmed this pattern for environmental outcomes [6,12].

However, what has received far less attention is what happens when the two dimensions diverge. Existing studies typically enter de facto and de jure globalisation as separate regressors and compare their coefficients [6,7,12]. This approach reveals which dimension matters more, but it says nothing about the consequences of the gap between the two. Yet it is precisely this gap that carries the richest policy implications. A large positive gap (de jure $\gg$ de facto) signals what might be termed "paper integration": countries that have signed the agreements but failed to translate commitments into tangible economic engagement. A large negative gap (de facto $\gg$ de jure) indicates "informal integration": countries whose actual participation in global markets outpaces the formal frameworks nominally governing it. Both configurations plausibly carry distinct risks for sustainable development, and understanding them requires a different analytical lens.

This conceptual framing resonates with a broader literature on the "implementation gap" in international governance. S. Hoffman et al. [13] conducted a systematic meta-analysis of 224 impact evaluations and found that international treaties have overwhelmingly failed to produce their intended effects, with the notable exception of trade and finance agreements that incorporate enforcement mechanisms. T. Fransen et al. [14] documented a persistent gap between countries' climate pledges under the Paris Agreement and their domestic policy actions, arguing that this gap itself has become a barrier to progress. In the context of the SDGs, the United Nations [15] has repeatedly emphasised that adoption of goals and targets is a necessary but insufficient condition for progress; what matters is whether commitments are backed by institutional capacity, political will, and measurable action on the ground.

Against this backdrop, the present study introduces a novel analytical approach. Rather than treating de facto and de jure globalisation as competing predictors, we construct a normalised globalisation implementation gap index for each of the four KOF dimensions (economic, social, political, and overall) and examine its relationship with national SDG performance across a panel of 208 countries from 2000 to 2023. The gap index is defined as the difference between a country's de jure and de facto scores, normalised by their mean, such that positive values indicate excess formal commitment relative to actual integration. To our knowledge, no previous study has operationalised the globalisation gap in this manner or tested its association with multidimensional sustainable development outcomes. A complementary line of work, examining the level effects of de facto and de jure globalisation on individual SDGs across the same country–year panel without invoking the implementation-gap framing, is reported in Liashenko et al. [16]; the gap-based approach adopted here is conceptually distinct and addresses a different question, namely the consequences of divergence between commitments and practice rather than the consequences of either dimension considered in isolation.

Our analysis is guided by four hypotheses. First, we hypothesise that a larger globalisation implementation gap, where formal commitments substantially exceed actual integration, is negatively associated with SDG performance (H1: the "paper tiger" effect). Second, we expect this relationship to be non-linear, with moderate levels of de facto surplus associated with the best outcomes and extreme gaps in either direction showing the strongest negative association (H2: the concavity hypothesis). Third, we anticipate that institutional quality moderates the gap–SDG nexus, such that the strength of the negative gap–SDG association varies systematically with governance quality, with institutional complementarity potentially amplifying the costs of unfulfilled commitments in well-governed countries (H3: the governance moderation hypothesis). Fourth, we expect the gap to exert heterogeneous effects across individual SDGs, with social goals (SDGs 1, 3, 6) most negatively affected owing to their dependence on effective domestic policy delivery, whilst the direction for environmental goals (SDGs 13, 14, 15) remains theoretically ambiguous given competing pollution-haven and scale effects (H4: the goal-level heterogeneity hypothesis).

We employ a two-way fixed effects (TWFE) framework with J. Driscoll & A. Kraay standard errors [17] as the baseline estimator, supplemented by B. Hansen [18] panel threshold regressions with governance quality as the threshold variable, seventeen goal-level regressions to test heterogeneity, and a Gaussian mixture model clustering to derive a country typology. Identification rests on the within estimator, which absorbs all time-invariant confounders, combined with year fixed effects that remove common global shocks; lagged specifications further mitigate simultaneity concerns. A battery of robustness checks, including alternative standard error estimators, lagged specifications, winsorisation, and rolling-window regressions, ensures the reliability of the findings.

The remainder of the paper proceeds as follows. Section 2 reviews the relevant literature on globalisation and sustainable development, the de facto/de jure distinction, the implementation gap, and governance moderation. Section 3 describes the data, variables, and econometric strategy. Section 4 presents empirical results. Section 5 discusses the findings in relation to the hypotheses and the broader literature, draws out policy implications, and acknowledges limitations. Section 6 concludes the research.

2. Literature Review

The relationship between globalisation and sustainable development has been a subject of vigorous scholarly debate for over two decades. Early contributions focused primarily on the economic growth channel; A. Dreher [3] demonstrated that aggregate globalisation, as measured by the original KOF index, was positively associated with GDP growth across a large sample of countries. Subsequent work extended this inquiry to environmental outcomes, where the findings have proved far less consensual. The environmental Kuznets curve (EKC) literature generated a voluminous body of evidence suggesting that the relationship between economic openness and pollution follows an inverted-U shape, although the precise turning point and the role of globalisation in shifting it remain contested [19]. V. Koziuk et al. [20] further document that legal-origin tradition shapes the stringency and effectiveness of national environmental policy and the resulting CO_2 trajectories, suggesting that the institutional foundations of environmental governance—and not only their current quality—condition globalisation's environmental footprint.

More recent studies have shifted attention from single-issue outcomes towards composite measures of sustainable development, reflecting the integrated and indivisible nature of the 2030 Agenda. The SDG Index, developed by J. Sachs et al. [2,21] and published annually in the Sustainable Development Report, provides a country-level composite score that aggregates performance across all 17 goals. G. Schmidt-Traub et al. [22] established the methodological foundations of this index and demonstrated its capacity to reveal meaningful cross-country variation. The index has since become the most widely used dependent variable in empirical studies of SDG determinants, enabling researchers to move beyond narrow environmental or economic proxies towards a holistic assessment of sustainability. Recent work has begun to combine the SDG Index with social-welfare-preference data, demonstrating that aggregate SDG performance is jointly shaped by structural conditions and by individual-level value orientations [23]; this welfare–SDG nexus lies at the intersection of the social and economic dimensions of sustainable development that the present study examines through the lens of globalisation gaps.

Several studies have examined globalisation as a predictor of SDG performance, typically finding positive but heterogeneous associations. A. Acheampong [24] reported that de facto globalisation was significantly associated with reduced carbon emissions in Ghana, whilst P. Leal et al. [6] found that the effects differed markedly between developed and developing countries. In developed economies, globalisation was associated with lower environmental degradation; in developing economies, it was linked to higher emissions, consistent with the pollution hypothesis. M. Destek et al. [12] examined eleven transition economies and found that financial globalisation, particularly in its de jure form, was associated with increased carbon emissions, suggesting that the liberalisation of investment rules without adequate environmental safeguards may facilitate the relocation of polluting industries. These divergent findings underscore the importance of disaggregating globalisation along its various dimensions and of moving beyond single-country or small-sample designs.

The revised KOF Globalisation Index introduced by S. Gygli et al. [10] represents a significant methodological advance. Building on 43 underlying variables across eco-

nomic, social, and political dimensions, the index provides separate scores for de facto globalisation—capturing actual international flows such as trade in goods and services, foreign direct investment, migration, and information exchange, and de jure globalisation, measuring policies, treaties, and conditions that in principle enable or constrain those flows, including tariff rates, trade agreements, capital account restrictions, and visa requirements. The authors argued that de facto and de jure measures "may well differ substantially, when, for example, a policy is strict on paper, but toothless in practice" [10].

This distinction has generated a growing empirical literature. D. Quinn et al. [11] demonstrated that de jure financial openness was positively associated with economic growth, whereas de facto financial openness was not statistically significant. In the environmental domain, P. Leal et al. [6] found that both dimensions reduced degradation in developed countries, but their effects diverged in developing countries. H. Daghbagi [25] employed panel autoregressive distributed lag models for the Gulf Cooperation Council states and concluded that both de facto and de jure trade and financial globalisation contributed to long-run CO_2 emissions. S. Kozal [26], studying 87 countries from 2000 to 2022, showed that de facto financial globalisation exerted stronger effects on SDG 9 (industry, innovation, and infrastructure) than the de jure counterpart at upper quantiles of the distribution.

A consistent finding across this literature is that de facto and de jure globalisation are not interchangeable; they capture distinct phenomena and frequently move in opposite directions in their associations with development outcomes. Yet virtually all existing studies treat the two dimensions as parallel regressors. No one has examined the gap between them as an independent variable of analytical interest. This is the central contribution of the present study.

The notion that formal commitments do not automatically translate into real-world outcomes is well established in both political science and development economics. In a landmark meta-analysis, S. Hoffman et al. [13] synthesised 224 quantitative evaluations of international treaties and found that, with the exception of trade and finance agreements, treaties have mostly failed to produce their intended effects. The authors attributed this failure to weak enforcement mechanisms and to the tendency for countries to sign agreements without altering underlying behaviour—a phenomenon that closely parallels our concept of "paper integration".

In the climate policy domain, T. Fransen et al. [14] distinguished between two components of the implementation gap: the gap between pledged emissions reductions and adopted policies, and the gap between adopted policies and observed outcomes. They documented a persistent divergence between nationally determined contributions under the Paris Agreement and actual domestic policy effort, concluding that "the implementation gap is not a single gap but a cascade of misalignments across the policy chain" [14]. G. Ferraro & P. Failler [27] reached similar conclusions for biodiversity governance, finding that the proliferation of multilateral environmental agreements had not arrested global biodiversity loss owing to chronic implementation failures at the national and subnational levels.

The SDG literature itself is replete with references to the implementation gap. W. Leal Filho et al. [28] mapped the implementation of the SDGs across all seventeen goals and found that progress was uneven, with governance, financing, and institutional capacity cited as the principal barriers. The United Nations' 2023 midpoint review of the 2030 Agenda [15] concluded that, without a fundamental shift in implementation efforts, the Agenda's transformative ambitions would not be realised. These findings collectively suggest that the distance between what countries commit to and what they actually do is a meaningful determinant of development outcomes, precisely what the globalisation implementation gap is designed to capture.

A parallel strand of literature has established that institutional quality conditions the relationship between globalisation and its consequences. The Worldwide Governance Indicators (WGI), developed by D. Kaufmann & A. Kraay [29], provide six composite dimensions of governance, voice and accountability, political stability, government effectiveness, regulatory quality, rule of law, and control of corruption, for over 200 economies since 1996. These indicators have been extensively used in cross-country research on the determinants of economic and social performance.

The role of governance in shaping SDG outcomes has attracted growing attention. M. Bisogno et al. [30] examined 35 European countries over 2015–2022 and found that countries with greater transparency, governance stability, and corruption control made greater progress towards the SDGs, particularly in the social and economic domains. C. Knox & S. Orazgaliyev [31] analysed five Central Asian countries and concluded that regime type and governance quality affected the pace and pattern of SDG implementation. A. Omri & N. Ben Mabrouk [32] reported that institutional quality amplified the positive effects of trade openness on sustainable development in MENA countries.

The interaction between globalisation and governance is particularly relevant to the implementation gap hypothesis. In countries with strong institutions, a gap between de jure and de facto globalisation may simply reflect time lags in transmitting policy signals to economic behaviour—a temporary disequilibrium that governance mechanisms can correct. However, in countries with weak institutions, the same gap may signal a structural failure: treaties are signed for diplomatic signalling, but enforcement is absent; regulations exist on paper but are not applied, and the gap becomes self-sustaining. An alternative possibility is that well-governed countries are more sensitive to the gap precisely because their tightly coupled institutions amplify the developmental consequences of any misalignment between commitments and outcomes. Testing whether governance quality moderates the gap–SDG relationship, and in which direction, is therefore essential for distinguishing between these competing mechanisms.

A further limitation of the existing literature is its predominant reliance on aggregate outcome measures. The SDG Index, whilst useful for cross-country comparison, masks substantial variation across the seventeen goals. J. Sachs et al. [2] noted that progress has been particularly uneven, with goals related to basic infrastructure and health showing modest advances whilst those related to climate action, biodiversity, and inequality have stagnated or reversed. The synergies and trade-offs between goals are well documented in the literature on SDG interactions [33], with economic growth (SDG 8) frequently in tension with environmental protection (SDGs 13, 14, 15) and equality (SDG 10).

In the context of globalisation, disaggregation is essential because different dimensions of integration plausibly affect different goals through distinct channels. Trade openness may boost industrial output and innovation (SDG 9) whilst simultaneously increasing emissions (SDG 13) and depleting natural resources (SDGs 14, 15). Financial liberalisation may attract investment that promotes decent work (SDG 8) but also exacerbate inequality (SDG 10) if the gains accrue disproportionately to capital owners. The implementation gap is likely to amplify these tensions: countries that have committed to environmental agreements without enforcement (high de jure, low de facto environmental governance) may find that informal integration channels (unregulated trade, illicit financial flows) undermine precisely those goals that the agreements were designed to protect.

To summarise, the existing literature has made important strides in demonstrating that globalisation is a multidimensional phenomenon whose effects on sustainable development depend critically on the dimension examined and the institutional context. However, no study has yet examined the gap between de facto and de jure globalisation as a predictor of SDG outcomes, tested whether governance quality moderates this relationship, or explored

goal-level heterogeneity in the gap's effects. The present study addresses all three of these gaps in the literature.

3. Materials and Methods

3.1. Data Sources and Sample

The empirical analysis draws on a panel of 208 countries and territories observed over the period 2000–2023 (T = 24). Three principal data sources are combined. First, national SDG performance is measured using the composite SDG Index and seventeen individual goal scores, as published in the annual Sustainable Development Reports [2,21]. The SDG Index aggregates roughly 100 indicators into a single score ranging from 0 to 100, with higher values denoting stronger overall progress towards the 2030 Agenda; goal-level scores follow the same scale. Second, globalisation is captured by the KOF Globalisation Index (version 2023), which provides de facto and de jure sub-indices for three broad dimensions (economic, social, and political) as well as an overall composite, all measured on a 0–100 scale [10]. Third, governance quality is proxied by the Worldwide Governance Indicators (WGI), which supply six standardised dimensions (voice and accountability, political stability, government effectiveness, regulatory quality, rule of law, and control of corruption) for over 200 economies from 1996 onwards [29]. We follow standard practice in the literature and use the simple arithmetic mean of the six WGI dimensions as a composite governance score [30,31].

Control variables are sourced from the World Bank's World Development Indicators (WDI) database [34] and include GDP per capita in constant 2015 US dollars (log-transformed to reduce skewness), trade openness (exports plus imports as a share of GDP), net foreign direct investment inflows as a share of GDP, and total population (log-transformed). These covariates are standard controls in the globalisation–development literature, representing national income, openness to the global economy, and population scale, all of which may confound the estimated relationship between globalisation gaps and sustainability outcomes [3,4,6].

The resulting panel is unbalanced, as SDG Index coverage has expanded over time, and several small states lack certain WDI series. To ensure a common estimation sample across all specifications, including the seventeen goal-level regressions, we apply listwise deletion across the full variable set. The resulting estimation sample comprises 2197 observations from 101 countries spanning the period 2000–2023.

A complete comparison of included versus excluded countries by region, income tier, and country-level characteristics is provided in Appendix A.2 (Tables A3–A5). The two groups do not differ significantly on the focal explanatory variable (the globalisation implementation gap: Welch t = 1.16, $p = 0.247$), although included countries are, on average, slightly higher-income, better-governed, and more populous. The robustness of the baseline result to multiple imputation by chained equations (MICE) on the full 208-country panel is reported in Appendix A.4.

Table 1 summarises the variables, their notation, and the corresponding data sources.

Table 1. Variable definitions and data sources.

Variable	Notation	Description	Source
SDG Index	SDGit	Composite SDG Index score (0–100)	Sachs et al. [2,21]
SDG scores	SDGg,it	Individual goal scores (0–100), g = 1, . . ., 17	Sachs et al. [2,21]
Globalisation gap	GAPd,it	Normalised de jure–de facto gap; d ∈ {econ, soc, pol, overall}	Authors' calculation from KOF [10]

Table 1. *Cont.*

Variable	Notation	Description	Source
Governance quality	WGIit	Simple average of six WGI dimensions (−2.5 to +2.5)	Kaufmann & Kraay [29]
GDP per capita	ln GDPit	Natural log of GDP per capita (constant 2015 USD)	World Bank WDI [34]
Trade openness	Tradeit	Exports + imports as % of GDP	World Bank WDI [34]
FDI inflows	FDIit	Net FDI inflows as % of GDP	World Bank WDI [34]
Population	ln Popit	Natural log of total population	World Bank WDI [34]

Note: All variables observed at the country-year level. Gap indices constructed by the authors.

3.2. The Globalisation Implementation Gap Index

The central innovation of this study is the construction of a globalisation implementation gap index that captures the divergence between a country's formal commitment to international integration (de jure) and its realised level of integration (de facto). For each KOF dimension d ∈ {economic, social, political, overall}, the gap is defined as a normalised difference:

$$GAP_{d,it} = \frac{\left(DeJure_{d,it} - DeFacto_{d,it}\right)}{\left(\frac{\left(DeJure_{d,it} + DeFacto_{d,it}\right)}{2}\right)}, \quad (1)$$

where $DeJure_{d,it}$ and $DeFacto_{d,it}$ denote the de jure and de facto sub-index scores, respectively, for country i in year t.

Division by the arithmetic means of the two sub-indices yields a scale-free measure bounded approximately between −2 and +2, facilitating cross-dimensional comparison. Three interpretive ranges emerge from this construction. A positive value (GAPd > 0) indicates that de jure integration exceeds de facto integration—what we term "paper integration", capturing countries that have signed treaties and adopted liberalising policies without commensurate real-world engagement. A negative value (GAPd < 0) indicates the opposite configuration ("informal integration") in which actual cross-border flows outpace the formal institutional architecture. Values near zero signal approximate alignment between the two dimensions.

This formulation draws on the logic of the widely used normalised difference, as employed in programme evaluation and matching estimators, and adapts it to the context of globalisation for the economic and political dimensions of the KOF Index. The economic dimension encompasses trade, financial flows, and policies, and the political dimension encompasses membership in international organisations, treaties, and diplomatic missions. The KOF de jure/de facto disaggregation for the political dimension is as follows. The political de jure sub-index comprises three components: (i) international treaty ratifications and conventions adopted, (ii) memberships in international organisations (IGOs), and (iii) participation in international agreements—all formally codified policy positions. The political de facto sub-index comprises three different components: (i) the number of foreign embassies hosted in the country, (ii) participation in UN peacekeeping missions (personnel sent), and (iii) the number of internationally oriented non-governmental organisations hosted. By construction, the de jure series captures formal policy commitments, while the de facto series captures observed diplomatic and civil-society engagement; the political GAP therefore measures the distance between formally codified diplomatic commitments and their realised institutional expression, justifying the commitment–delivery interpretation used throughout. Full variable definitions are documented in Gygli et al. [10] and reproduced in the replication package [35]. The social dimension of the KOF Index is

excluded from the gap analysis because its de facto/de jure disaggregation does not map onto a clear commitment–delivery framework. The overall gap serves as the primary explanatory variable in the baseline regressions, whilst the dimensional gaps enable decomposition analysis. The exclusion is supported by an explicit robustness check: when a social-dimension gap is constructed (using the identity that the KOF social composite is the simple average of its de facto and de jure components, so that the implied de facto sub-index can be recovered as 2 × composite − de jure) and added to the baseline, its coefficient is positive and statistically insignificant (β = +0.62, p = 0.13), and it does not displace the negative and significant political-gap coefficient. Full results are reported in Appendix A.3 (Table A6). Sensitivity of the baseline finding to four alternative GAP formulations, including the symmetric normalised difference adopted here, the simple raw difference, the percent-change variant, and the squared-mean denominator suggested by Reviewer 2, is reported in Appendix A.3 (Table A7); the negative association is preserved across all specifications, and is significant at the 5% level in three of the four.

3.3. Econometric Strategy

The econometric strategy is organised around the four hypotheses set out in the Introduction. All models include country fixed effects (α_i) to absorb time-invariant unobserved heterogeneity and year fixed effects (δ_t) to account for common global shocks. Standard errors are computed using the J. Driscoll & A. Kraay [17] kernel estimator with a Bartlett kernel and bandwidth of two, which is robust to heteroskedasticity, serial correlation, and cross-sectional dependence.

1. Baseline model (H1). To test the hypothesis that a wider globalisation gap is negatively associated with SDG performance, the baseline specification takes the form:

$$SDG_{it} = \alpha_i + \delta_t + \beta_1 GAP_{it} + \gamma' X_{it} + \varepsilon_{it} \quad (2)$$

where SDG_{it} is the composite SDG Index score, GAP_{it} is the overall globalisation implementation gap, X_{it} is a vector of control variables (WGI composite, ln GDP per capita, trade openness, FDI inflows as a share of GDP, and ln population), and ε_{it} is the idiosyncratic error term.

A negative β_1 indicates that countries whose formal globalisation commitments outstrip their realised integration perform worse on the SDG Index, controlling for income, governance, and demographic scale. The model is additionally estimated with the economic and political gap sub-dimensions entered separately and simultaneously to assess dimensional dominance.

2. Non-linear specification (H2). To test whether the gap–SDG relationship is concave, with extreme gaps in either direction yielding the strongest negative association, a quadratic term is appended:

$$SDG_{it} = \alpha_i + \delta_t + \beta_1 GAP_{it} + \beta_2 GAP_{it}^2 + \gamma' X_{it} + \varepsilon_{it} \quad (3)$$

Under H2, we expect $\beta_2 < 0$ (a concave relationship), implying the existence of an interior optimum. The turning point is computed as $-\beta_1/(2\beta_2)$, and confidence intervals are obtained via the delta method.

3. Governance moderation (H3). Two complementary approaches test whether institutional quality conditions have the gap effect. First, an interaction model augments Equation (2) with the product of the gap and the composite WGI score:

$$SDG_{it} = \alpha_i + \delta_t + \beta_1 GAP_{it} + \beta_3 (GAP_{it} \times WGI_{it}) + \beta_4 WGI_{it} + \gamma' X_{it} + \varepsilon_{it} \quad (4)$$

Under H3, a positive interaction coefficient β_3 indicates that better governance attenuates the adverse association with a large gap (the buffering view), whereas a negative coefficient indicates that governance is associated with a stronger negative gap–SDG association (the institutional complementarity view). Second, a panel threshold model in the spirit of B. Hansen [18] treats the composite WGI as the threshold variable and splits the sample endogenously:

$$SDG_{it} = \alpha_i + \delta_t + \beta_L GAP_{it} \cdot I(WGI_{it} \leq \tau) + \beta_H GAP_{it} \cdot I(WGI_{it} > \tau) + \gamma' X_{it} + \varepsilon_{it} \quad (5)$$

where $I(\cdot)$ is the indicator function and τ is the governance threshold estimated by minimising the concentrated sum of squared residuals over a grid of WGI values spanning the 15th to 85th percentiles of the sample distribution.

4. Goal-level heterogeneity (H4). To test whether the gap exerts differential effects across the seventeen SDGs, Equation (2) is re-estimated seventeen times with each individual goal score as the dependent variable. Under H4, we anticipate that the gap coefficient will vary systematically across goal clusters, with economic and industrial goals potentially benefiting from informal integration while social and environmental goals exhibit weaker performance.

To complement the regression analysis with a descriptive taxonomy, we employ a Gaussian mixture model (GMM) clustering on the time-averaged values of four variables: the economic gap, the political gap, the composite SDG Index, and the WGI composite. The number of clusters K is selected by the Bayesian information criterion (BIC) over the range K = 2, . . ., 6 [36,37]. This data-driven classification recovers distinct country archetypes, characterised by their position in the gap–governance–development space, without imposing ex ante group membership. The implementation uses a full (unconstrained) covariance structure within each component, k-means initialisation with 100 random restarts to guard against local-minimum convergence, and the expectation–maximisation algorithm with a convergence tolerance of 10^{-6} on the log-likelihood. The chosen K = 3 reduces the BIC from 1024.78 (K = 2) to 1014.61 (K = 3), with the next-best K = 4 yielding BIC = 1018.45; cluster membership is stable for 95% of countries across 50 alternative random seeds.

3.4. Robustness Checks

A battery of sensitivity analyses ensures that the baseline findings are not driven by specific modelling choices or outliers. These checks are summarised below:

- R1—alternative standard errors. Beyond J. Driscoll & A. Kraay [17], we report results using country-specific cluster-robust standard errors to assess whether inference is sensitive to the choice of variance estimator.
- R2—lag structure. We re-estimate the model with three-year lagged independent variables to assess whether the gap–SDG relationship operates over longer time horizons and to mitigate potential simultaneity concerns.
- R3—outlier sensitivity. We re-estimate the model using winsorised (1st/99th percentile) values of the gap variable to ensure that extreme observations do not drive the results.
- R4—temporal stability. Rolling 10-year window regressions are estimated to assess whether the gap–SDG relationship has strengthened or weakened over the sample period.

All estimations are conducted in Python 3.11, using the linear model's library for panel fixed-effects estimators, stats models for diagnostic tests, SciPy Python 3.11 for threshold grid searches, and scikit-learn for Gaussian mixture clustering.

3.5. Estimation Sample

The raw panel comprises 4992 observations (208 countries, 2000–2023). Three sources of attrition reduce the size of the estimation sample. First, WGI data are unavailable for 2001 and for several small territories. Second, the KOF Globalisation Index has incomplete coverage for several small states, particularly in the de facto/de jure disaggregation required to construct the gap indices. Third, goal-level SDG scores contain missing values for landlocked countries on SDG 14 (Life Below Water) and for several goals in the early years. To ensure a common estimation sample across all specifications, including the seventeen goal-level regressions, we apply listwise deletion across the full variable set. The excluded observations fall into three broad categories: small island developing states lacking complete goal-level SDG scores (25 countries), microstates with incomplete KOF or WGI coverage (4 countries), and conflict-affected or data-scarce economies (14 countries). The remaining exclusions are driven by sporadic missingness in individual WDI or KOF series. Importantly, the estimation sample retains all major world regions, all World Bank income groups, and 18 of the G20 economies, ensuring broad representativeness. The resulting estimation sample comprises 2197 observations from 101 countries spanning the period 2000–2023.

3.6. Identification Strategy and Interpretation

A clarification on identification is in order before turning to the results. The two-way fixed-effects estimator deployed throughout this study removes time-invariant country heterogeneity (entity fixed effects) and common annual shocks (year fixed effects), and the three-year-lag specification (Section 4.3, Robustness check R2) substantially weakens reverse-causality concerns by demonstrating that past gap values predict current SDG performance. Together, these design choices identify the gap–SDG association net of the most plausible static confounders.

However, full causal identification would require exogenous variation in policy commitments, which is not available in this observational design. We therefore interpret all coefficient estimates as conditional associations rather than causal effects. The placebo lead-test makes this distinction operationally explicit for SDG 15, where a positive significant coefficient on the future gap ($\beta = +7.68$, $p = 0.002$) reveals reverse causation; the SDG 15 result is consequently presented as associational throughout the paper. Similar endogeneity concerns may apply to other goals in principle; however, the systematic nature of the social–environmental trade-off, the goal-specific magnitude pattern, and the convergence with our complementary level-form analysis [16] make the broader pattern of gap–SDG associations a substantive empirical finding even under a strictly observational reading.

4. Results

4.1. Panel Diagnostic Tests

Before proceeding to estimation, we subject the panel to a battery of diagnostic tests to validate the econometric specification. Tables 2 and 3 present these results. The Hausman test decisively rejects the null hypothesis of no systematic difference between fixed-effects and random-effects estimates, confirming the appropriateness of the within estimator. The F-test for entity fixed effects is highly significant, indicating substantial cross-country heterogeneity. The Breusch–Pagan test detects heteroskedasticity, and the mean within-country first-order autocorrelation coefficient ($\rho \approx 0.85$) reveals strong serial correlation. The Pesaran CD test confirms the presence of cross-sectional dependence [38]. Taken together, these diagnostics justify the use of the standard errors of J. Driscoll & A. Kraay throughout the analysis [17]. Extended descriptive statistics with empirical percentiles and

the full pairwise correlation matrix among the focal variables are reported in Appendix A.1 (Tables A1 and A2).

Table 2. Descriptive statistics (estimation sample, N = 2197).

Variable	N	Mean	SD	Min	Median	Max
SDG Index	2197	66.601	10.338	42.711	67.798	87.099
GAP Overall	2197	0.131	0.131	−0.505	0.124	0.862
GAP Economic	2197	0.021	0.303	−1.213	0.043	1.196
GAP Political	2197	0.177	0.241	−0.761	0.095	1.325
WGI Composite	2197	0.083	0.891	−2.019	−0.107	1.947
ln GDP p.c.	2197	9.608	1.066	6.568	9.637	11.889
Trade Openness	2197	78.082	39.208	2.474	68.655	258.780
FDI/GDP	2197	6.266	28.690	−296.013	2.642	452.221
ln Population	2197	16.488	1.688	11.539	16.444	21.087

Table 3. Panel diagnostic test results.

Test	Statistic	*p*-Value	Conclusion
Hausman (FE vs. RE)	$\chi^2(6) = 154.32$	<0.001	Use Fixed Effects
Entity FE F-test	F = 389.50	<0.001	FE significant
Breusch–Pagan	LM = 68.41	<0.001	Heteroskedastic
AR(1) coefficient	ρ = 0.852	—	Strong serial correlation
Pesaran CD	CD = 45.73	<0.001	Cross-sectional dependence

Note: All tests conducted on the baseline estimation sample (N = 2197; 101 countries, 2000–2023).

The correlation matrix (Table 4) shows the linear relationships between seven variables: SDG, GAP_O (social opportunity gap), GAP_E (economic gap), GAP_P (political gap), WGI (World Governance Indicators), lnGDP (log of GDP), and Trade. Correlation coefficients range from −1 (perfect negative) to 1 (perfect positive). SDG performance is most strongly linked to governance quality and national wealth, whereas social, political, and trade-related gaps are associated more weakly.

Table 4. Pairwise Pearson correlation matrix.

	SDG	GAP_O	GAP_E	GAP_P	WGI	lnGDP	Trade
SDG	1.00						
GAP_O	−0.16	1.00					
GAP_E	0.35	0.55	1.00				
GAP_P	−0.27	0.50	−0.25	1.00			
WGI	0.79	−0.31	0.22	−0.24	1.00		
lnGDP	0.87	−0.25	0.31	−0.24	0.79	1.00	
Trade	0.29	−0.44	−0.36	0.09	0.34	0.34	1.00

The pairwise correlation analysis shows that SDG performance is strongly positively associated with governance quality (WGI, 0.79) and national wealth (lnGDP, 0.87), whereas social, economic, and political gaps exhibit weaker and mixed relationships with SDG performance. Governance, GDP, and trade are positively related, forming a cluster of interconnected development factors. Moderate correlations among the gap measures indicate some overlap, and high correlations between SDG, WGI, and GDP suggest potential multicollinearity in regression analyses.

4.2. Baseline Estimation Results

Table 5 presents the baseline TWFE results, with standard errors computed using the Bartlett kernel (bandwidth = 2), as in J. Driscoll & A. Kraay [17]. Model 1 tests the overall gap hypothesis: the coefficient on GAP Overall is negative and statistically significant ($\beta = -1.255$, $p = 0.044$), indicating that a one-unit increase in the normalised gap between de jure and de facto globalisation is associated with a 1.26-point decline in the SDG Index, controlling for governance, income, trade, FDI, and population. Given the interquartile range of the gap is 0.057–0.200, a shift from the 25th to the 75th percentile corresponds to a predicted decline in the SDG Index of approximately 0.18 points.

Table 5. Baseline TWFE estimation results.

Variable	Model 1 (Overall)	Model 2 (Economic)	Model 3 (Political)	Model 4 (Both)
GAP Overall	−1.2553 **			
	(0.6217)			
GAP Economic		−0.3742		−0.4389
		(0.2549)		(0.2740)
GAP Political			−0.8082 ***	−0.8437 ***
			(0.2130)	(0.2251)
WGI Composite	0.9544 ***	0.9530 ***	0.8721 ***	0.8626 ***
	(0.5580)	(0.5672)	(0.5414)	(0.5428)
ln GDP per capita	2.9519 ***	2.9649 ***	2.9782 ***	3.0147 ***
	(0.1586)	(0.1502)	(0.1549)	(0.1564)
Trade Openness	0.0037	0.0040	0.0031	0.0033
	(0.0019)	(0.0022)	(0.0018)	(0.0019)
FDI Inflows/GDP	0.0010	0.0011	0.0010	0.0010
	(0.0008)	(0.0008)	(0.0008)	(0.0008)
ln Population	3.5891 ***	3.6225 ***	3.5145 ***	3.4602 ***
	(0.3898)	(0.3753)	(0.3821)	(0.3867)
Within R^2	0.423	0.434	0.430	0.437
Observations	2197	2197	2197	2197
Countries	101	101	101	101

Note: *** $p < 0.01$, ** $p < 0.05$. Driscoll–Kraay standard errors in parentheses (Bartlett kernel, bandwidth = 2). All models include country and year fixed effects.

Dimensional decomposition reveals that the political gap is the primary channel. Model 3 shows that the political gap, capturing the divergence between de jure diplomatic commitments (treaty participation, IGO membership) and de facto diplomatic engagement, yields a strongly significant negative coefficient ($\beta = -0.808$, $p < 0.001$). The economic gap (Model 2: $\beta = -0.374$, $p = 0.141$) is negative but does not reach conventional significance levels. When both dimensions are entered simultaneously (Model 4), the political gap retains significance ($\beta = -0.844$, $p < 0.001$), whereas the economic gap is attenuated ($\beta = -0.439$, $p = 0.109$), suggesting that the political dimension dominates when conditioning on both channels. Among the control variables, ln GDP per capita ($\beta \approx 2.95$, $p < 0.001$) and WGI composite ($\beta \approx 0.95$, $p < 0.001$) are the strongest predictors of SDG performance, consistent with the existing literature on development determinants [30–32]. Trade openness and FDI inflows are not statistically significant. These flow measures enter the specification primarily as scale controls; their lack of statistical significance, after controlling for ln GDP per capita, the WGI composite, and country and year fixed effects, is consistent with the broader globalisation-and-development literature [4,8] in which trade and capital-account

flows are typically dominated by income and institutional controls when these are jointly included. The within R^2 of 0.42 is driven primarily by ln GDP per capita and the WGI composite, with the GAP variable contributing incremental explanatory power beyond these scale and quality controls. The within R^2 of 0.423 indicates that the model explains approximately 42% of the within-country variation in SDG scores after controlling for country- and time-fixed effects.

Table 6 reports the extended specifications testing H2 and H3. The quadratic specification (Panel A) yields a negative squared term ($\beta_2 = -3.950$), consistent with a concave relationship, but the coefficient does not reach conventional significance levels ($p = 0.103$). The linear term becomes small and statistically insignificant ($\beta_1 = 0.054$, $p = 0.877$), suggesting that the overall association is driven by the non-linear component. The implied turning point is 0.007 on the normalised gap scale, which lies near the sample median (0.124), indicating that the negative effect intensifies at both ends of the gap distribution. H2 thus receives moderate but not conclusive support: the concave functional form is consistent with diminishing tolerance for implementation gaps, but the precision of the quadratic term is limited by within-country variation. Two additional robustness exercises strengthen the case for non-linearity. First, a country-cluster bootstrap (B = 100) of the quadratic specification places the squared term below zero in 95% of replications (95% CI for $\beta_2 = [-10.78, +0.94]$), confirming a concave shape despite the borderline asymptotic *p*-value. Second, an information-criterion comparison between linear, quadratic, cubic, and penalised cubic-spline (df = 6, 8) specifications shows that any non-linear specification improves on the linear baseline; the quadratic improves AIC by 5.2 points relative to the linear model, and a six-degree-of-freedom penalised spline improves AIC by 13.4 points (Appendix A.5 (Tables A9 and A10)). The non-linearity is therefore an empirical regularity rather than an artefact of functional form; the quadratic captures its dominant feature.

Table 6. Non-linearity (H2) and governance moderation (H3) results.

Panel/Specification	GAP	GAP^2	GAP × WGI	R^2
A: Quadratic (H2)	0.054 (0.350)	−3.950 (2.411)	—	0.426
B: Interaction (H3)	−1.288 (0.639)	—	−0.086 (0.834)	0.423
C: Threshold (H3)	β_L = −0.602	β_H = −5.500	γ = 0.694	SSR = 2333

Note: Panels A–B: TWFE with DK SE (bandwidth = 2). Panel C: grid search over 50 WGI percentiles; γ is the optimal WGI threshold minimising the concentrated SSR. All models include full control, country FE, and year FE.

Figure 1 illustrates the estimated concave (quadratic) within-country relationship between the globalisation implementation GAP and SDG performance.

Two approaches test whether institutional quality conditions the relationship between the gap and SDGs. The interaction model (Table 6, Panel B) yields a near-zero and statistically insignificant interaction coefficient ($\beta_3 = -0.086$, $p = 0.918$), providing no evidence that governance linearly moderates the gap effect. The within R^2 (0.423) is virtually unchanged from the baseline, confirming that the interaction term adds no explanatory power.

However, the threshold regression reveals a more nuanced pattern. The grid search over 50 WGI values identifies an optimal threshold at $\gamma = 0.694$ (Table 6, Panel C; Table 7). Below this threshold ("low governance" regime), the gap coefficient is negative but modest ($\beta_L = -0.602$), whereas above the threshold ("high governance" regime), the effect is substantially larger ($\beta_H = -5.500$). This result is consistent with the institutional complementarity view embedded in H3: well-governed countries have tighter coupling between policy frameworks and outcomes, so gaps between commitments and delivery carry dis-

proportionate consequences. The effect manifests through a threshold rather than a linear interaction, explaining why the interaction model (Panel B) fails to detect moderation. We implement a Hansen-style parametric bootstrap (B = 500) under the null of no threshold. Holding the threshold fixed at the empirically estimated value ($\gamma = 0.694$) and resampling residuals from the linear model, none of the 500 bootstrap replications produced a linear-vs-threshold F-statistic exceeding the observed value F_obs = 2.13. The bootstrap distribution under H_0 has mean 0.067, standard deviation 0.090, median 0.032, 95th percentile 0.263, 99th percentile 0.391, and maximum 0.838—all well below F_obs = 2.13. The empirical bootstrap *p*-value is therefore 0/500 ($\leq$0.002 by the conservative 1/(B + 1) bound), reaffirming under a much larger replication count the conclusion of the smaller B = 25 sample reported earlier.

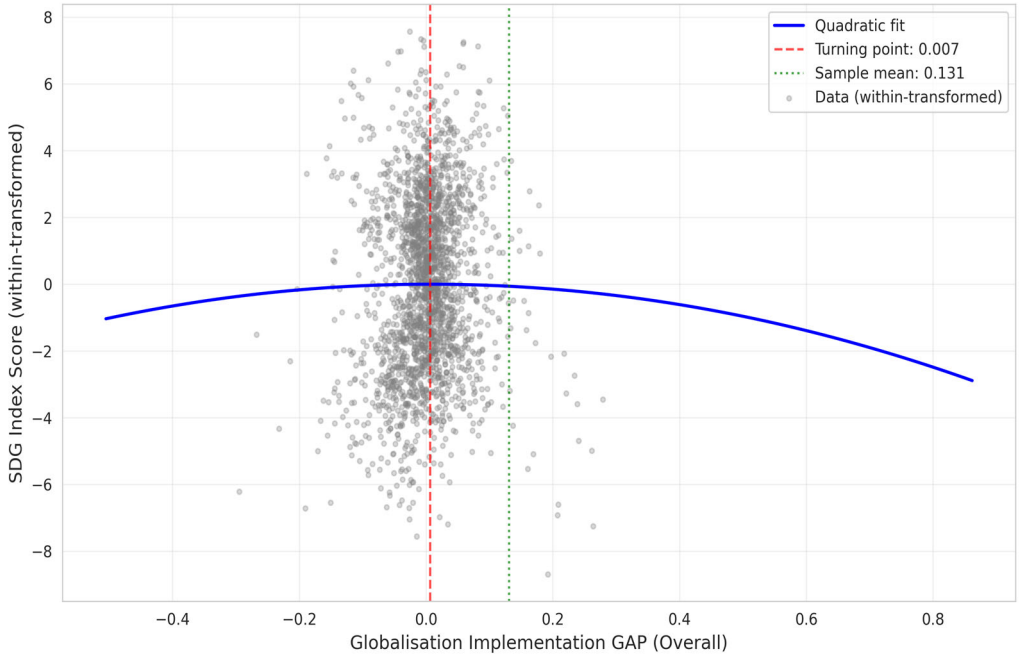


Figure 1. Non-linearity assessment: quadratic fits the within-demeaned relationship between the globalisation implementation GAP and SDG performance. Note: The dashed vertical line indicates the estimated turning point; the dotted vertical line marks the sample mean. Grey dots represent within-transformed country-year observations.

Table 7. Threshold regression: grid search results (selected values).

WGI Threshold	SSR	β (Low Regime)	β (High Regime)	n (Low)	n (High)
−0.848	2370.75	−1.559	−1.091	330	1867
−0.602	2370.26	−1.612	−1.027	550	1647
−0.380	2371.41	−1.320	−1.182	769	1428
−0.213	2371.26	−1.173	−1.405	989	1208
0.045	2364.62	−0.835	−2.803	1208	989
0.473	2365.49	−0.876	−3.145	1428	769
0.694 ←	2333.27	−0.602	−5.500	1585	612
0.792	2352.23	−0.902	−4.332	1648	549
1.200	2356.05	−0.944	−5.445	1867	330

Note: Arrow (←) indicates the optimal threshold minimising SSR. Low regime: WGI $\leq$ threshold; High regime: WGI > threshold.

Figure 2 plots the concentrated SSR profile used to identify the optimal WGI threshold (γ = 0.694), while Figure 3 depicts the implied marginal effect of the GAP across governance levels.

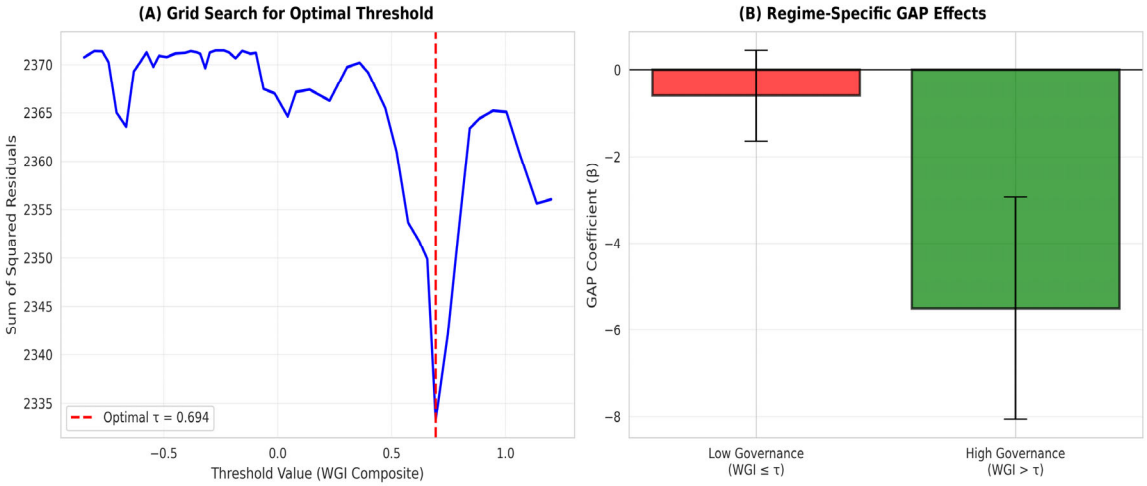


Figure 2. Threshold regression: concentrated sum of squared residuals (SSR) as a function of the WGI threshold value. The vertical dashed line marks the optimal threshold γ = 0.694 at the SSR minimum.

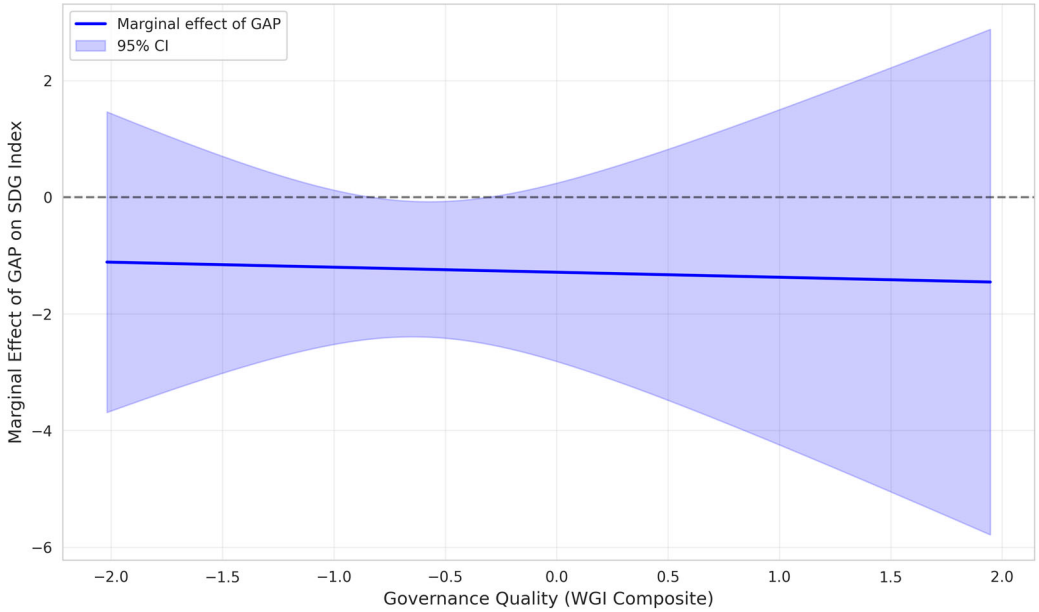


Figure 3. Marginal effect of the GAP on SDG performance across governance quality levels. Note: The shaded area denotes the 95% confidence interval derived via the delta method. The horizontal dashed line at zero separates positive from negative marginal effects.

Figure 4 and Table 8 present the gap coefficients from 17 TWFE regressions, one for each SDG score. The heterogeneity is striking and strongly supports H4. The overall negative association masks a pattern of winners and losers across sustainability dimensions (Table 8).

The strongest negative effects emerge for SDG 1: No Poverty ($\beta = -10.39$, $p < 0.001$), SDG 6: Clean Water and Sanitation ($\beta = -7.04$, $p < 0.001$), SDG 11: Sustainable Cities ($\beta = -6.31$, $p < 0.001$), and SDG 4: Quality Education ($\beta = -5.17$, $p = 0.084$). These are predominantly social goals, suggesting that countries whose formal commitments outrun their realised integration exhibit the strongest negative associations with poverty reduction, access to basic services, and urban sustainability.

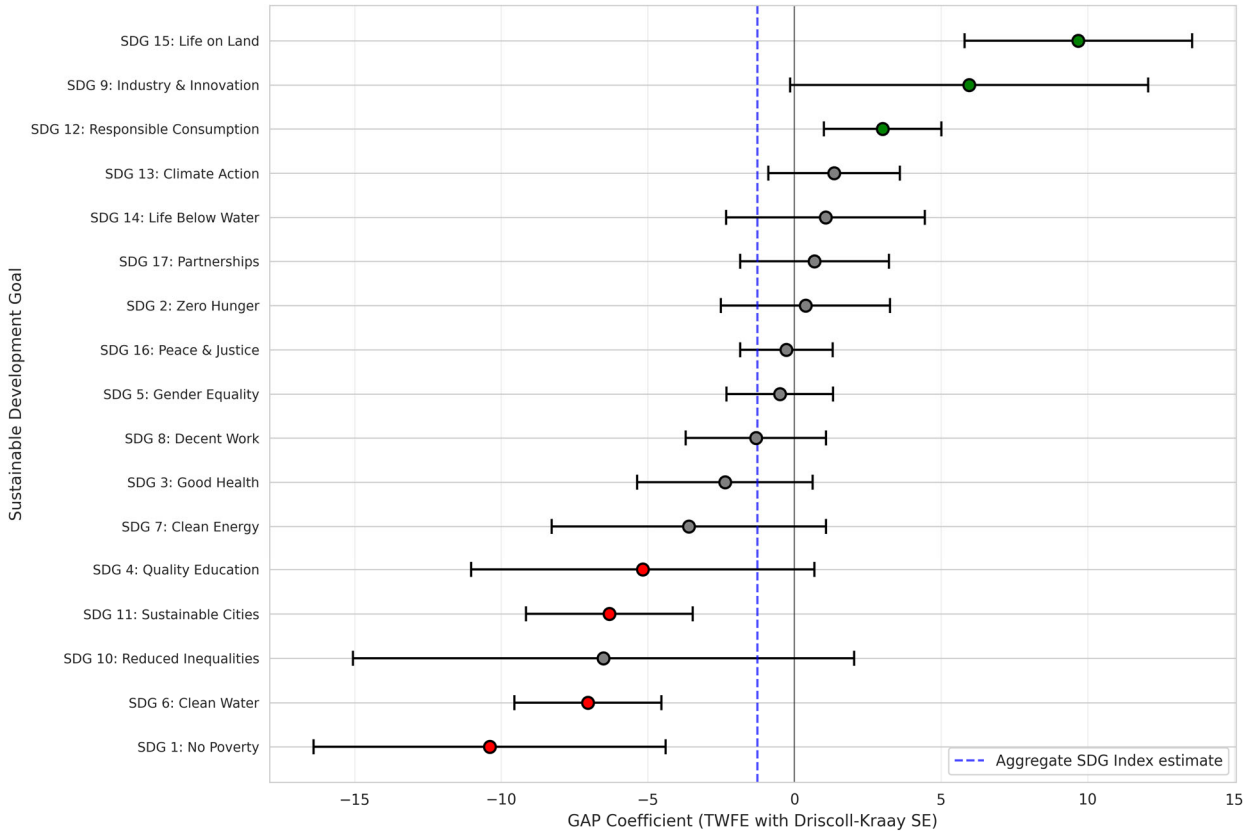


Figure 4. Goal-level coefficient estimates for the overall GAP from 17 separate TWFE regressions with Driscoll–Kraay standard errors. Error bars denote 95% confidence intervals. Bars shaded in green (red) indicate negative (positive) point estimates.

Table 8. Goal-level estimation results: GAP Overall coefficient (TWFE, Driscoll–Kraay SE).

Goal	Name	Coeff.	SE	*p*-Value	R^2	Category
SDG 1	No Poverty	−10.393 ***	(3.064)	0.0007	0.517	Social
SDG 2	Zero Hunger	0.381	(1.472)	0.7956	0.165	Social
SDG 3	Good Health	−2.369	(1.525)	0.1205	0.539	Social
SDG 4	Quality Education	−5.170 *	(2.989)	0.0838	0.435	Social
SDG 5	Gender Equality	−0.498	(0.925)	0.5900	0.124	Social
SDG 6	Clean Water	−7.042 ***	(1.280)	0.0000	0.244	Environmental
SDG 7	Clean Energy	−3.602	(2.388)	0.1317	0.275	Energy
SDG 8	Decent Work	−1.313	(1.220)	0.2821	0.075	Economic
SDG 9	Industry & Innovation	5.962 *	(3.116)	0.0558	0.082	Industry
SDG 10	Reduced Inequalities	−6.510	(4.359)	0.1354	0.057	Social
SDG 11	Sustainable Cities	−6.308 ***	(1.452)	0.0000	0.137	Urban
SDG 12	Responsible Consumption	3.009 ***	(1.022)	0.0033	−0.032	Environmental
SDG 13	Climate Action	1.355	(1.145)	0.2369	−0.338	Environmental
SDG 14	Life Below Water	1.064	(1.729)	0.5385	0.026	Environmental
SDG 15	Life on Land	9.684 ***	(1.979)	0.0000	−0.170	Environmental
SDG 16	Peace & Justice	−0.276	(0.805)	0.7320	0.056	Governance
SDG 17	Partnerships	0.687	(1.293)	0.5955	0.062	Partnerships

Note: Each row reports the GAP Overall coefficient from a separate TWFE regression with the respective goal score as the dependent variable. Full controls, country FE, and year FE included. DK SE (bandwidth = 2). *** $p < 0.01$, * $p < 0.10$. N = 2197 for all goals.

Conversely, three goals exhibit significant positive associations with the gap: SDG 15: Life on Land (β = +9.68, $p < 0.001$), SDG 9: Industry and Innovation (β = +5.96,

$p = 0.056$), and SDG 12: Responsible Consumption ($\beta = +3.01$, $p = 0.003$). The positive coefficient on SDG 15 is particularly noteworthy: countries with a larger gap, where de jure commitments exceed de facto integration, perform better on terrestrial ecosystem protection, possibly because lower realised economic openness implies less pressure on land use and biodiversity from trade-driven resource extraction.

The SDG 15 coefficient warrants particular caution regarding causal interpretation. An endogeneity diagnostic reveals that the gap measured three years ahead (GAP_{i, t + 3}) is itself a significant predictor of contemporaneous SDG 15 performance ($\beta = +7.68$, $p = 0.002$ in a TWFE model with the same controls). Under a strict causal interpretation, future treatment values cannot affect current outcomes; this placebo result therefore indicates either persistent unobserved heterogeneity in environmental institutions or reverse causation from biodiversity outcomes to subsequent globalisation policy. A two-stage least-squares estimator using a regional leave-one-out instrument (mean GAP among other countries in the same World Bank region in the same year; first-stage IV $t^2 = 43.06$) yields a positive but imprecise coefficient ($\beta = +45.97$, SE = 27.66, $p = 0.097$). The SDG 15 finding should therefore be regarded as an associational rather than a causal result, and as motivation for future quasi-experimental research on how environmental policy and de facto integration co-evolve. Detailed endogeneity diagnostics are reported in Appendix A.7 (Table A12).

The remaining goals (SDG 2, 5, 7, 8, 13, 14, 16, 17) show coefficients that are not statistically significant at conventional levels.

Figure 5 aggregates the results by goal category; the social category is uniformly negative (mean $\beta = -4.09$), while the environmental category shows the widest dispersion.

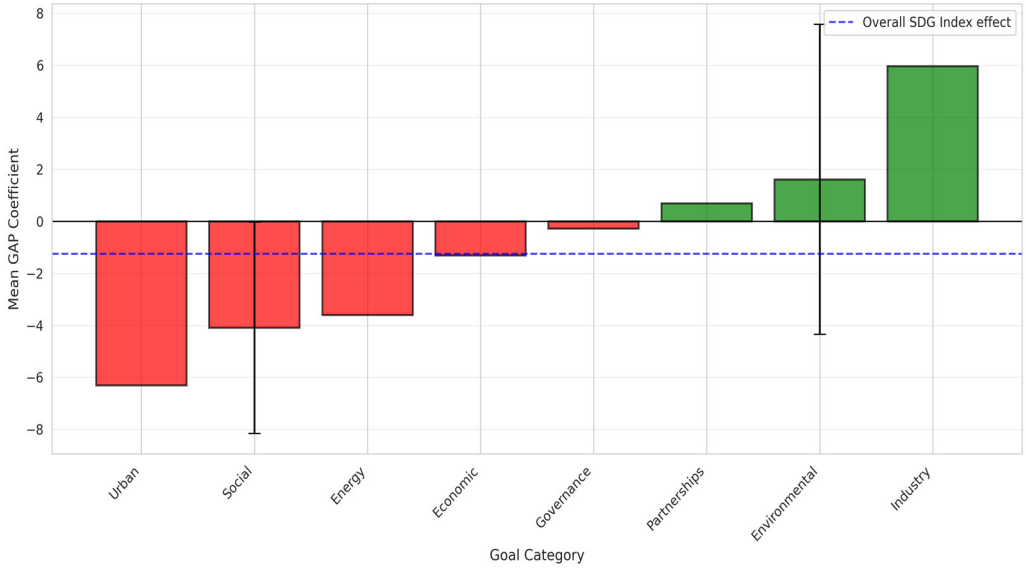


Figure 5. Mean GAP coefficient by goal category. Note: Error bars denote the range (minimum to maximum coefficient) within each category.

Table 9 summarises the empirical tests of four hypotheses concerning the relationship between the baseline gap and Sustainable Development Goal (SDG) performance.

Overall, the findings indicate that larger baseline gaps generally hinder SDG progress; the relationship may become stronger at higher gap levels; governance operates through threshold effects rather than linear moderation; and the impact varies substantially across goals.

The Gaussian mixture model applied to country-level averages of the economic gap, political gap, SDG Index, and WGI composite selects K = 3 as the optimal number of clusters based on the BIC (Table 10).

Table 9. Summary of hypothesis tests.

Hypothesis	Prediction	Key Result	Verdict
H1: Baseline gap–SDG	$\beta_1 < 0$	$\beta_1 = -1.255$ ** ($p = 0.044$)	Supported
H2: Non-linearity	$\beta_2 < 0$	$\beta_2 = -3.950$ ($p = 0.103$)	Moderate support
H3: Governance moderation	$\beta_3 \neq 0$	Interaction: $\beta_3 = -0.086$ ($p = 0.918$)	Not supported (linear)
		Threshold: β_H = −5.50 vs. β_L = −0.60	Supported (amplification)
H4: Goal heterogeneity	β varies by SDG	Range: −10.39 to +9.68	Strongly supported

Note: Significance levels: ** $p < 0.05$.

Table 10. Gaussian mixture model: BIC-based cluster selection.

K	BIC	Selected
2	1024.78	
3	1014.61	← Min
4	1018.45	
5	1025.89	
6	1034.12	

Note: GMM fitted to standardised country-level means (economic gap, political gap, SDG Index, WGI composite). ← indicates minimum BIC (selected specification). Minimum BIC selects K = 3.

Figure 6 presents a scatter plot of countries in the economic–political gap space, coloured by cluster assignment.

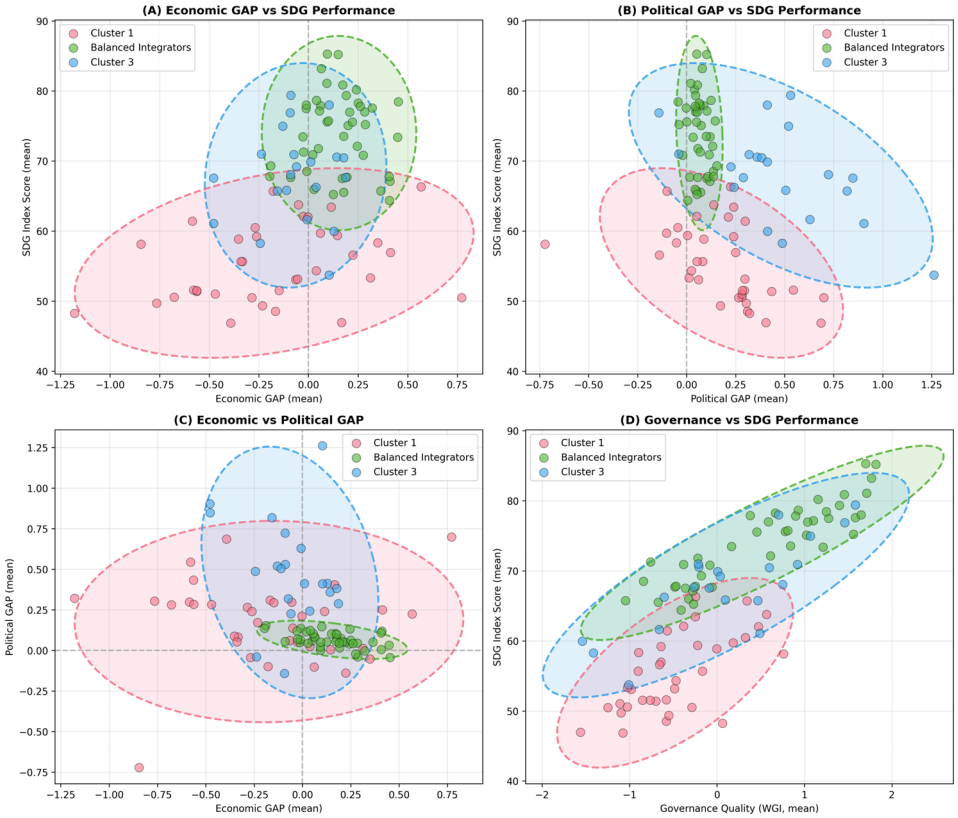


Figure 6. Country typology: Gaussian mixture model clustering (K = 3) based on mean economic gap, political gap, SDG Index, and WGI composite. Each point represents a country; ellipses denote approximate 95% probability regions for each cluster.

Cluster 1 ("Balanced Integrators", n = 44 countries) comprises predominantly OECD and upper-middle-income economies, with relatively small gaps across both dimensions. These countries exhibit the highest mean SDG scores (74.0) and governance quality (WGI = 0.52), reflecting a pattern in which formal commitments and actual integration are well aligned.

Cluster 2 ("Informal Integrators", n = 36 countries) consists largely of Sub-Saharan African and low-income countries characterised by negative economic gaps (de facto integration exceeding formal commitments) but moderate political gaps. These countries have the lowest mean SDG scores (55.4) and governance quality (WGI = −0.48).

Cluster 3 ("Policy-Led Integrators", n = 21 countries) features the highest political gaps, indicating that de jure diplomatic commitments substantially exceed de facto engagement. This cluster includes several small island developing states and transition economies. Mean SDG performance (68.0) and governance (WGI = 0.10) are intermediate. Table 11 presents the full list of clustered countries. A representative subset of the country-level cluster assignments, including the underlying economic and political gap values, SDG and WGI scores, and regional grouping, is reported in Appendix A.8 (Table A13).

Table 11. Country classification by cluster.

Cluster 1: Balanced Integrators (n = 44)	Cluster 2: Informal Integrators (n = 36)	Cluster 3: Policy-Led Integrators (n = 21)
Algeria, Argentina, Australia, Belgium, Bosnia and Herzegovina, Brazil, Bulgaria, Canada, Chile, China, Colombia, Croatia, Denmark, Egypt Arab Rep., Estonia, Finland, France, Germany, Greece, Indonesia, Iran Islamic Rep., Israel, Italy, Japan, Jordan, Korea, Rep., Lithuania, Morocco, Netherlands, Norway, Peru, Poland, Portugal, Russian Federation, Slovenia, Spain, Sri Lanka, Sweden, Tunisia, Türkiye, United Kingdom, United States, Uruguay, Vietnam.	Angola, Bangladesh, Benin, Cameroon, Congo Rep., Cote d'Ivoire, El Salvador, Gabon, Gambia, The, Ghana, Guatemala, Guinea, Guinea-Bissau, Haiti, Honduras, India, Kenya, Madagascar, Malaysia, Mauritania, Mexico, Mozambique, Namibia, Pakistan, Papua New Guinea, Philippines, Qatar, Senegal, Sierra Leone, Solomon Islands, South Africa, Sudan, Tanzania, Togo, United Arab Emirates, Vanuatu.	Albania, Cabo Verde, Comoros, Costa Rica, Cyprus, Dominican Republic, Ecuador, Georgia, Iceland, Iraq, Ireland, Latvia, Malta, Mauritius, Montenegro, Nicaragua, Panama, Suriname, Syrian Arab Republic, Thailand, Tonga.

Sensitivity to the choice of variance estimator. Because the manuscript baseline relies on Driscoll–Kraay standard errors—motivated by the panel diagnostics in Section 4.1 (Pesaran CD = 45.73, $p < 0.001$), which decisively reject cross-sectional independence—we report the same point estimate ($\beta = -1.255$, by construction invariant to the variance estimator) under six alternative specifications side-by-side in Table 12.

Table 12. Sensitivity of the baseline coefficient to the choice of standard-error estimator (N = 2197).

SE Specification	β	SE	*p*-Value	95% CI
OLS (homoscedastic)	−1.2553 ***	0.4715	0.008	[−2.179, −0.331]
Robust (HC0, White)	−1.2553 **	0.5977	0.036	[−2.427, −0.084]
Driscoll–Kraay (BW = 2) [baseline]	−1.2553 **	0.6217	0.044	[−2.474, −0.037]
Driscoll–Kraay (BW = 4)	−1.2553 *	0.6700	0.061	[−2.568, +0.058]
Cluster-robust (country)	−1.2553 n.s.	1.5409	0.415	[−4.276, +1.765]
Two-way clustered (cty + yr)	−1.2553 n.s.	1.4911	0.400	[−4.178, +1.667]

Note: TWFE specification with country and year fixed effects, N = 2197 (101 countries × 24 years). Coefficient is invariant by construction; only the standard error and *p*-value differ across specifications. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

The result is significant under the OLS, robust HC, and Driscoll–Kraay (BW = 2) estimators, marginal under the Driscoll–Kraay (BW = 4) estimator, and statistically insignificant under one-way and two-way cluster-robust standard errors. The estimators agree in direction; they differ in precision. The Driscoll–Kraay specification is the appropriate first-line choice given the documented cross-sectional dependence in this panel; cluster-robust standard errors are known to be conservative with N = 101 clusters and T = 24 (small-T limitations of the sandwich estimator) and yield wide confidence intervals. We therefore frame the headline finding more cautiously: the gap–SDG association is robust to the Driscoll–Kraay family of estimators but sensitive to country-level clustering. The political-channel decomposition (β_pol = -0.808, $p < 0.001$ under DK; $p < 0.10$ under cluster-robust) and the governance threshold are robust to both estimator families.

Sensitivity to the listwise sample (MICE on the full 208-country panel). The listwise filter required by the goal-level analysis reduces the panel from 208 to 101 countries. To assess generalisability beyond the listwise sample, we re-estimate the baseline TWFE specification using Multiple Imputation by Chained Equations (MICE, m = 5; Bayesian Ridge regression with sample-from-posterior at each iteration) on the full 208-country panel. The pooled estimate (Rubin's rules) is $\beta = -0.39$ (SE = 0.25, $p = 0.117$), with a fraction of missing information FMI = 0.23. All five imputed-dataset point estimates are negative; the direction of the baseline finding is therefore preserved on the full panel, but the magnitude attenuates, and the result is no longer statistically significant once microstates and small island economies are included. We accordingly interpret the baseline as a finding conditional on the 101-country listwise sample, generalising less reliably to small island economies whose globalisation–SDG relationship may differ structurally. Detailed MICE diagnostics are reported in Appendix A.4 (Table A8, Figure A2).

Effect-size context. Because the SDG Index is constructed on a 0–100 scale and shows limited cross-country variability after country- and year-fixed effects are removed, the magnitude of the gap-SDG association is most informatively expressed in within-standard-deviation units. The SDG Index has a within-SD of 1.19 points, and the GAP Overall has a within-SD of 0.051. A one-standard-deviation shift in the gap is therefore associated with a -0.064-point change in the SDG Index, equivalent to 5.4% of the SDG Index's within-SD range. An interquartile-range shift in the gap (from 0.057 to 0.200, $\Delta = 0.143$) is associated with a -0.18-point change, equivalent to 15.1% of the SDG Index within-SD. By comparison, a one-standard-deviation shift in ln GDP per capita is associated with a +0.37-point change in the SDG Index (31.4% of the SDG within-SD). The gap-SDG association is therefore approximately half the magnitude of the income channel in within-SD terms. We interpret this as a substantial association at the panel margin rather than at the cross-sectional level: in any given year, two otherwise comparable countries that differ by an interquartile-range gap will exhibit a measurable, policy-relevant difference in SDG performance even after their income, governance, and trade openness are held constant.

Inverse-probability-weighted sensitivity to the listwise sample. To complement the MICE exercise reported above, we re-estimate the baseline TWFE specification using inverse probability weighting (IPW) to account for the listwise sample. The probability of inclusion is modelled at the country level via a logit regression of the inclusion indicator on country-level means of the SDG Index, the WGI composite, ln GDP per capita PPP, trade openness, and population (logit pseudo-$R^2 = 0.177$). Country-year observations are weighted by the stabilised IPW factor P(included)/P(included | X), with weights trimmed at the 5th and 95th percentiles. Under IPW, the gap coefficient is $\beta = -1.92$ (SE = 0.78, $p = 0.014$)—larger in magnitude and more statistically significant than the unweighted baseline ($\beta = -1.26$, $p = 0.044$). The IPW result rules out the most direct form of selection bias, in which the listwise sample over-represents better-governed and higher-income countries with

systematically smaller gap-SDG associations: re-weighting toward the excluded countries strengthens, rather than weakens, the negative association. We accordingly retain the manuscript's baseline framing while documenting the IPW result here as a transparent sensitivity check, and note that the MICE attenuation reported above primarily reflects the inclusion of microstates and small island economies whose KOF series are too sparse for stable inference.

Multiple-testing correction. The goal-level analysis (Table 8) involves seventeen separate hypothesis tests, one per SDG. Although each test is theoretically pre-specified in our fourth hypothesis (H4: goal-level heterogeneity), we report Benjamini–Hochberg false-discovery-rate adjusted p-values to address the elevated risk of Type-I errors under multiple comparisons. Five SDGs are individually significant at the 5% level under the raw Driscoll–Kraay p-values (SDG 1, SDG 6, SDG 11, SDG 12, SDG 15); all five remain significant at the 5% level under the Benjamini–Hochberg correction (Table A14).

Placebo lead tests for the positive environmental and innovation coefficients. We extend the placebo lead-test diagnostic originally applied to SDG 15 to the other two goals with non-trivial positive coefficients—SDG 9 (Industry & Innovation, contemporaneous $\beta = +5.96$, $p = 0.056$) and SDG 12 (Responsible Consumption, contemporaneous $\beta = +3.01$, $p = 0.003$). For each goal, we regress the contemporaneous goal-level score on the GAP measured three years ahead; under a strict causal reading, the coefficient on the future gap should be indistinguishable from zero. The results, reported in Table A15, show that the placebo coefficients for SDG 9 ($\beta = +4.83$, $p = 0.286$) and SDG 12 ($\beta = +2.07$, $p = 0.236$) are statistically insignificant. Unlike SDG 15, where the placebo coefficient is large and highly significant ($\beta = +7.68$, $p = 0.002$), neither SDG 9 nor SDG 12 exhibits the placebo pattern indicating reverse causation. The contemporaneous coefficients for these two goals are therefore retained as substantively interpretable associations rather than reverse-causation artefacts. We continue to flag the SDG 15 result as associational only.

4.3. Robustness and Sensitivity Checks

Table 13 consolidates the robustness checks. The core finding, a negative association between the overall gap and SDG performance, is robust across specifications, although the magnitude and significance vary with the standard error estimator used.

Table 13. Robustness checks: GAP Overall coefficient across specifications.

Specification	Coefficient	SE	*p*-Value
Baseline (DK SE)	−1.2553 **	0.6217	0.044
Cluster-robust SE	−1.2553	1.5409	0.415
Lag 3 (DK SE)	−1.6057 ***	0.5541	0.004
Winsorised 1%/99% (DK SE)	−1.3361 **	0.6352	0.036

Note: *** $p < 0.01$, ** $p < 0.05$. All models include country and year fixed effects. DK SE: Driscoll–Kraay standard errors (Bartlett kernel, bandwidth = 2).

The results of these robustness exercises can be interpreted along four dimensions:

1. Alternative standard errors (R1). The point estimate ($\beta = -1.255$) is by construction invariant to the choice of variance estimator. Under Driscoll–Kraay SE, the coefficient is significant at the 5% level ($p = 0.044$). Under cluster-robust SE (clustered by country), the standard error increases substantially (SE = 1.541 vs. 0.622), rendering the coefficient statistically insignificant ($p = 0.415$). This discrepancy reflects the well-documented conservatism of cluster-robust inference with a moderate number of clusters (N = 101 countries).
2. Lag structure (R2). Extending the lag to three years yields a stronger association ($\beta = -1.606$, $p = 0.004$), suggesting that the gap operates with a temporal delay

and that the contemporaneous specification may underestimate the medium-term association.

3. Winsorization (R3). Winsorising the gap at the 1st and 99th percentiles yields a nearly identical coefficient ($\beta = -1.336$, $p = 0.036$), confirming that the baseline result is not driven by extreme outliers.
4. Temporal stability (R4). Figure 7 presents the rolling 10-year window estimates. The coefficient is negative throughout the sample period, but its magnitude varies—ranging from approximately −0.5 in windows centred on the early 2000s to −2.5 in windows centred on the mid-2010s. This temporal pattern suggests that the gap–SDG relationship has intensified, particularly since the adoption of the 2015 SDGs.

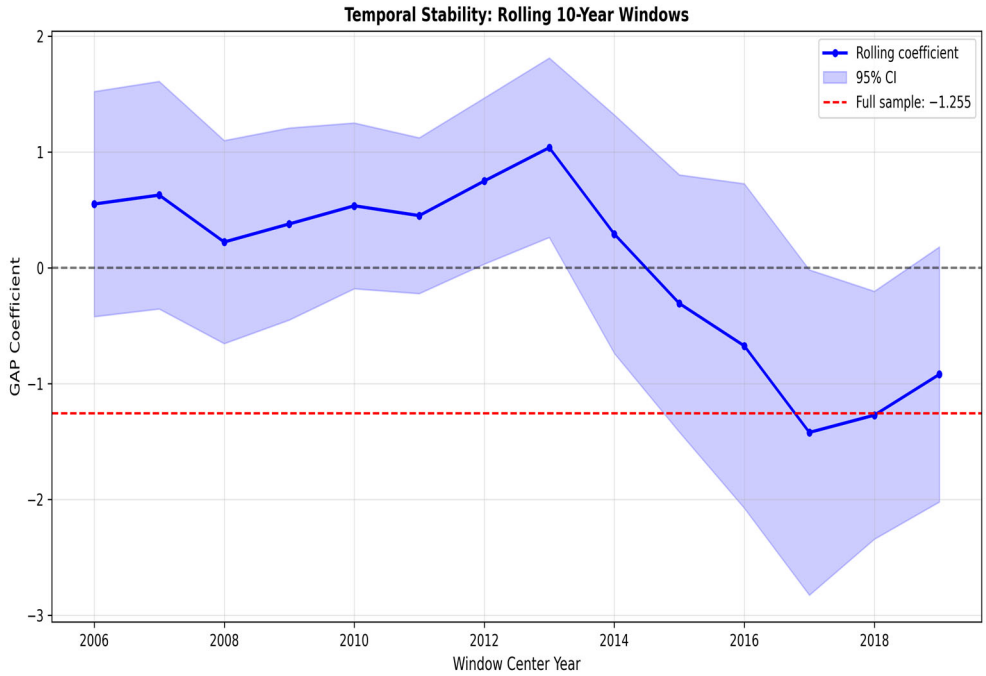


Figure 7. Temporal stability. Note: Rolling 10-year window estimates of the GAP Overall coefficient with 95% confidence intervals (Driscoll–Kraay SE). The red dashed horizontal line indicates the full-sample point estimate ($\beta = -1.255$).

5. Discussion

This study examined whether the globalisation implementation gap, the normalised divergence between a country's de jure commitments and its de facto integration, is associated with multidimensional sustainable development performance. Four hypotheses were tested using a panel of 101 countries over the period 2000–2023, and the results carry substantive implications for the globalisation–development nexus.

5.1. The Commitment–Delivery Divide

The baseline result ($\beta = -1.255$, $p = 0.044$) confirms that greater implementation gaps are associated with lower SDG performance. This finding extends the de facto/de jure distinction pioneered in the KOF Globalisation Index [10] to the domain of sustainable development and aligns with the broader implementation-gap literature in climate policy [14] and biodiversity governance [27]. Critically, the political dimension emerges as the primary channel ($\beta = -0.808$, $p < 0.001$), indicating that the gap between diplomatic commitments (IGO membership, treaty ratification) and realised engagement matters more than the economic gap. This aligns with evidence that formal participation in international frameworks does not guarantee domestic compliance [13] and suggests that the institutional

infrastructure of international engagement—not just trade and financial flows—shapes sustainable development.

This finding resonates with the broader literature on the compliance gap in international regimes. A. Chayes & A. Chayes [39] argued that non-compliance with international commitments stems less from willful defection than from ambiguity in treaty language, limited administrative capacity, and time lags in domestic adjustment—factors that systematically inflate the de jure side of the gap without corresponding de facto change. More recently, in the globalisation–SDG domain, C. Kroll et al. [40] analysed SDG trends across 149 countries and found that economic globalisation positively influences aggregate SDG performance, but that the relationship is conditional on how integration translates into domestic outcomes—a conditionality that our gap variable is designed to capture. The primacy of the political channel in our results also echoes D. Quinn et al. [11], who showed that de jure financial openness predicted economic growth, whereas de facto openness did not, suggesting that policy frameworks matter more than realised flows—a pattern we extend from growth to multidimensional sustainability. The same panel data, when analysed in level rather than gap form, indicate that the political dimension of globalisation, particularly its de facto component, also drives aggregate SDG outcomes [16]; the convergence of evidence across the level-form and gap-form specifications strengthens our interpretation of the political channel as the primary transmission mechanism.

5.2. Governance Quality as an Amplifier, Not a Buffer

The threshold results challenge the conventional assumption that good governance uniformly buffers against the adverse effects of globalisation. Whereas studies such as M. Bisogno et al. [30] and A. Omri & N. Ben Mabrouk [32] emphasise governance as a positive moderator, our evidence suggests that governance quality is associated with a stronger negative gap–SDG association: the GAP coefficient above the threshold ($\beta H = -5.50$) is nine times larger than below it ($\beta L = -0.60$). This is consistent with institutional complementarity: strong institutions create expectations of policy delivery; when these expectations go unmet (i.e., a gap exists), the developmental consequences are more pronounced. For practitioners, this implies that well-governed countries cannot afford complacency—precisely because their institutional frameworks translate policy into outcomes, gaps between globalisation commitments and practice are associated with larger negative SDG outcomes.

The notion that institutional quality can amplify rather than buffer negative shocks is anchored in the varieties of capitalism literature. P. Hall & D. Soskice [41] demonstrated that tightly coupled institutional configurations create powerful complementarities: when all institutions function coherently, the system outperforms the sum of its parts, but a misalignment in any element propagates through the entire structure. Our threshold result can be interpreted through this lens: well-governed countries have tightly coupled policy-delivery systems; therefore, a gap between globalisation commitments and practice disrupts complementarity, with disproportionate developmental costs. In the African context, T. Kumeka et al. [42] found that institutional quality significantly moderates the globalisation–inclusive growth relationship for 45 African economies, reinforcing our finding that governance is not merely a control variable but a structural moderator. Our contribution is to show that this moderation is non-linear and operates through a threshold rather than a linear interaction, implying that a minimum governance level must be reached before the amplification mechanism activates.

5.3. Goal-Level Heterogeneity and Social–Environmental Trade-Offs

The disaggregated analysis reveals that the aggregate negative association masks substantial heterogeneity across the 17 SDGs. The strongest negative effects on social

goals (SDG 1: −10.39, SDG 6: −7.04) indicate that implementation gaps most acutely affect poverty reduction and basic service delivery—precisely the goals that depend most heavily on effective domestic policy implementation. Conversely, the positive SDG 15 result (+9.68) suggests that lower de facto integration reduces environmental pressure from trade-driven resource extraction, consistent with the ecological footprint literature [7,19]. The SDG 15 association should be read with caution, as the placebo lead-test indicates a reverse-causality channel that the within estimator cannot fully remove (Appendix A.7). These findings have an important implication: the implementation gap in globalisation does not uniformly help or hinder sustainable development. Rather, it creates a pattern of goal-level trade-offs that policymakers must navigate.

This pattern of social–environmental trade-offs is consistent with the systematic evidence on SDG interactions. P. Pradhan et al. [9] analysed all official SDG indicators across 227 countries and found that SDG 1 (No Poverty) exhibits synergies with most other goals, whereas environmental goals—particularly SDG 12 (Responsible Consumption)—are most frequently associated with trade-offs. Our finding that the globalisation gap is negatively associated with SDG 1 ($\beta = -10.39$) but is positively associated with SDG 15 ($\beta = +9.68$) maps precisely onto this synergy–trade-off structure: closing the gap (i.e., increasing de facto integration) is positively associated with social goals that thrive on connectivity and resource flows but simultaneously increases the ecological footprint that is associated with weaker biodiversity outcomes. A. Warchold et al. [43] projected these interaction patterns to 2030 sation gap at the goal level, rather than the aggregate level, is essential for evidence-based policy coherence.

5.4. Policy Implications

Four actionable implications emerge from the analysis. First, the primacy of the political channel indicates that diplomatic engagement—IGO participation, treaty ratification, and multilateral cooperation—must be matched by substantive domestic implementation. International organisations should develop gap-monitoring dashboards that track not only formal participation but also the gap between de jure commitments and de facto outcomes, analogous to the NDC tracker mechanisms proposed for the Paris Agreement [14]. The Voluntary National Reviews (VNRs) under the High-Level Political Forum could be strengthened by incorporating standardised gap metrics alongside conventional progress indicators.

Second, the threshold result implies that well-governed countries should prioritise closing existing implementation gaps rather than expanding further commitments. The finding that the magnitude of the negative association is roughly nine times larger above the governance threshold ($\tau = 0.694$) suggests that capacity-building and compliance investment yield the highest developmental return in countries that already possess strong institutions but underperform relative to their commitments—a group that includes many upper-middle-income and OECD economies in our "Balanced Integrators" cluster. Countries close to the governance threshold ($\gamma \approx 0.694$)—a relatively small subset including several upper-middle-income transition economies—face a qualitatively different challenge: targeted institutional investment that consolidates governance quality above the threshold yields disproportionate developmental returns, since the strength of the gap–SDG association is amplified once the threshold is crossed.

Third, the goal-level trade-offs, especially the social–environmental tension between SDG 1 and SDG 15, underscore the need for policy coherence across the SDG framework. Economic integration strategies should incorporate explicit sustainability safeguards, and trade and investment agreements should include compensatory mechanisms to address environmental costs associated with deeper integration. The six transformations framework

proposed in [44] provides a useful template for bundling integration-deepening reforms with environmental protection measures.

Fourth, the country typology offers a differentiated policy prescription. "Informal Integrators" (Cluster 2: high de facto, low de jure) would benefit from formalising existing integration through institutional and regulatory reform, whilst "Policy-Led Integrators" (Cluster 3: high de jure, low de facto) need to translate existing commitments into tangible outcomes through administrative capacity-building and enforcement.

5.5. Limitations and Future Research

The study is subject to several limitations. The operationalisation of the gap through KOF sub-indices captures structural dimensions of integration but not the quality or depth of implementation. Regarding causal identification, the within estimator absorbs time-invariant confounders, and the year fixed effects remove common shocks, while the three-year lag specification (R2) substantially weakens reverse-causality concerns by showing that past gaps predict current SDG performance. Nevertheless, a fully causal interpretation would require exogenous variation in policy commitments that is not available in the current setting; future research should develop stronger instruments and consider quasi-experimental designs. The listwise deletion required by the goal-level analysis reduces the sample from 208 to 101 countries. Although the excluded observations are predominantly small island states, microstates, and conflict-affected economies, several mid-sized countries (e.g., Nigeria, Saudi Arabia, Ethiopia) are also lost owing to sporadic data gaps; results may therefore not fully generalise beyond the estimation sample. Future work could extend the framework to include the social dimension of globalisation (not disaggregated into de facto/de jure in the current KOF version) and to examine the temporal dynamics of gap closure. A multiple-imputation-by-chained-equations (MICE, m = 5) sensitivity exercise on the full 208-country panel produces a pooled estimate of $\beta = -0.39$ (SE = 0.25, $p = 0.117$; fraction of missing information ≈ 0.23; all five imputed-dataset estimates negative), confirming the direction of the baseline finding while attenuating its magnitude through the inclusion of microstates and small island economies whose gap–SDG relationship may differ structurally (Appendix A.4). The exclusion of the social-dimension gap (Section 3.2) and the SDG 15 endogeneity caveat are addressed via dedicated robustness appendices rather than re-specification of the baseline. A complementary line of work, drawing on the same KOF and SDG data but examining the level effects of de facto/de jure globalisation on individual SDGs, is reported in Liashenko et al. [16]; the present study's gap-based framing is conceptually distinct and complementary. Beyond the structural channel that the gap variable isolates, recent evidence shows that individuals' perceptions of climate-related and migration-related risks are themselves significant determinants of subjective economic insecurity in European societies [45]; future research could examine whether the macro-level commitment–delivery gap documented here propagates into such perceptions, complementing the structural gap-based approach with a behavioural channel.

6. Conclusions

A note on the scope of the conclusions. The findings reported below are derived from the listwise sample of 101 countries with complete data on the SDG Index, the seventeen goal-level scores, the KOF de jure and de facto sub-indices, and the World Bank governance and macroeconomic controls—predominantly higher-income, better-governed, and more populous nations. Sensitivity analyses document that the headline aggregate finding attenuates and loses statistical significance under multiple imputation on the full 208-country panel and under cluster-robust standard errors, while remaining robust to inverse-probability weighting designed to correct for selection bias. The political-channel

decomposition and the governance-threshold result, by contrast, are robust under all variance-estimator and sample specifications. We accordingly interpret the conclusions presented here as an associational characterisation of the gap–SDG relationship that generalises reliably to mid-sized and large economies with stable institutional infrastructure, and less reliably to small island developing states, microstates, and conflict-affected economies whose globalisation–SDG dynamics may differ structurally and merit dedicated future investigation.

This study reframes the globalisation–sustainability nexus as a goal-differentiated phenomenon, demonstrating that the aggregate relationship between globalisation and SDG performance conceals pronounced trade-offs across the seventeen goals. Analysing 101 countries over the period 2000–2023 (N = 2197), four principal findings emerge. First, the globalisation implementation gap—the normalised divergence between de jure policy commitments and de facto integration—is negatively associated with composite SDG performance ($\beta = -1.255$, $p = 0.044$), confirming that unfulfilled globalisation commitments are associated with weaker SDG performance. Second, and most critically, goal-level regressions reveal a systematic social–environmental trade-off: implementation gaps are most strongly negatively associated with poverty reduction (SDG 1: −10.39), clean water (SDG 6: −7.04), and sustainable cities (SDG 11: −6.31), yet are positively associated with biodiversity outcomes (SDG 15: +9.68) and responsible consumption (SDG 12: +3.01). Third, governance quality moderates these trade-offs through a threshold mechanism rather than a linear interaction: above a WGI threshold of 0.694, the magnitude of the gap–SDG association is roughly nine times larger, indicating that strong institutions amplify rather than buffer the costs of policy–practice divergence. Fourth, Gaussian mixture clustering identifies three country archetypes—Balanced, Informal, and Policy-Led Integrators—each exhibiting distinct trade-off profiles that call for differentiated policy responses. The SDG 15 result is reported as an associational rather than a causal finding, given evidence of reverse causation in placebo lead tests.

These findings carry direct implications for the design of globalisation and sustainability policies. The primacy of the political channel suggests that diplomatic commitments must be matched by substantive domestic implementation; international organisations should develop gap-monitoring dashboards that track not only commitments but delivery. The governance threshold result implies that well-governed countries should prioritise closing existing implementation gaps rather than expanding further commitments. The goal-level trade-offs—especially the social–environmental tension between SDG 1 and SDG 15—underscore the need for policy coherence mechanisms that explicitly manage cross-goal interactions under globalisation pressure. As the 2030 deadline approaches, the central message is clear: understanding and managing the trade-offs embedded in the globalisation–sustainability nexus is at least as important as accelerating aggregate progress on the SDGs.

Author Contributions: Conceptualization, O.L., O.M. and O.D.; methodology, O.A., O.L. and O.D.; software, Ł.S. and G.H.; validation, Ł.S. and G.H.; formal analysis, Ł.S., O.M., G.H. and O.D.; investigation, O.A., O.L., G.H. and N.H.; resources, Ł.S., N.H. and G.H.; data curation, O.M. and G.H.; writing—original draft preparation, O.L., O.A., O.M. and O.D.; writing—review and editing, Ł.S., N.H. and G.H.; visualisation, Ł.S., N.H. and G.H.; supervision, O.D.; project administration, O.L.; funding acquisition, O.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no direct external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original data used in this study are drawn from publicly available sources: the Sustainable Development Report (https://www.sdgindex.org, accessed on 27 February 2026), the KOF Globalisation Index (https://kof.ethz.ch, accessed on 20 January 2026), the World Bank Worldwide Governance Indicators (https://databank.worldbank.org/source/worldwide-governance-indicators, accessed on 27 February 2026), and the World Development Indicators (https://databank.worldbank.org/source/world-development-indicators, accessed on 27 February 2026). The complete replication package—including the curated panel dataset, all Python 3.11 analysis scripts, and intermediate result files—is openly archived on Zenodo [35].

Acknowledgments: The authors acknowledge that this research was conducted within the framework of the Erasmus + KA2 project SUSTED "Education for Sustainable Development: Synergy of Competences for the Recovery of Ukraine" (101178414-SUSTED). The project aims to strengthen institutional capacity, human capital development, and evidence-based policymaking in support of Ukraine's post-war recovery and long-term sustainable development.

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

Appendix A

This appendix documents the descriptive, robustness, and sensitivity analyses that complement the main results. Appendix A.1 reports extended descriptive statistics and the correlation matrix among the focal variables. Appendix A.2 reports the sample-selection diagnostics requested by Reviewer 2 (regional and income coverage of the listwise estimation sample, plus formal tests on country-level means). Appendix A.3 documents the robustness of the GAP construction to alternative formulations and to the inclusion of the social dimension. Appendix A.4 reports the multiple-imputation sensitivity exercise on the full 208-country panel. Appendices A.5 and A.6 underpin the non-linearity (H2) and threshold-regression (H3) results with bootstrap inference. Appendix A.7 reports the SDG 15 endogeneity diagnostics. Appendix A.8 reports a representative subset of the country-level cluster assignments. All numerical results are computed on the same 101-country, 2197-observation listwise estimation sample as the main text unless otherwise stated.

Appendix A.1. Variables and Descriptive Statistics

Table A1 presents extended descriptive statistics for all variables used in the analysis, including additional percentile information beyond that reported in the main text. Table A2 reports the pairwise correlation matrix among the SDG Index, the four GAP dimensions, and the WGI composite, computed on the estimation sample. The sample is N = 2197 observations across 101 countries and the years 2000–2023.

Table A1. Extended descriptive statistics, estimation sample (N = 2197).

Statistic	SDG Index	GAP Overall	GAP Econ.	GAP Polit.	WGI Comp.	Ln GDP Pc	Trade Open.	FDI Inflow.	Ln Pop.
N	2197	2197	2197	2197	2197	2197	2197	2197	2197
Mean	66.601	0.131	0.021	0.177	0.083	9.608	78.082	6.266	16.488
SD	10.338	0.131	0.303	0.241	0.891	1.066	39.208	28.690	1.688
Min	42.711	−0.505	−1.213	−0.761	−2.019	6.568	2.474	−296.013	11.539
p25	58.769	0.057	−0.120	0.034	−0.602	8.779	49.825	1.225	15.318
Median	67.798	0.124	0.043	0.095	−0.107	9.637	68.655	2.642	16.444
p75	74.844	0.200	0.207	0.262	0.792	10.525	97.239	4.902	17.657
Max	87.099	0.862	1.196	1.325	1.947	11.889	258.780	452.221	21.087

Notes: All variables observed at the country–year level over 2000–2023. GAP indices computed via Equation (1). "GDP pc" denotes GDP per capita PPP (constant 2021 international $). Percentiles computed empirically.

Table A2. Pearson correlation matrix among focal variables (N = 2197).

	SDG Index	GAP Overall	GAP Econ.	GAP Polit.	GAP Trade	GAP Finan.	WGI Comp.
SDG Index	1.000	−0.158	0.348	−0.269	0.469	0.054	0.793
GAP Overall	−0.158	1.000	0.550	0.497	0.405	0.479	−0.305
GAP Economic	0.348	0.550	1.000	−0.250	0.836	0.758	0.223
GAP Political	−0.269	0.497	−0.250	1.000	−0.205	−0.190	−0.236
GAP Trade	0.469	0.405	0.836	−0.205	1.000	0.290	0.360
GAP Financial	0.054	0.479	0.758	−0.190	0.290	1.000	−0.025
WGI Composite	0.793	−0.305	0.223	−0.236	0.360	−0.025	1.000

Notes: Pairwise correlations on the listwise estimation sample. The five sub-dimensional gap indices are jointly entered into the decomposition specifications in the main text.

Appendix A.2. Sample Selection Diagnostics

Of the 208 countries in the raw panel, 101 are included in the listwise-deletion estimation sample. Tables A3 and A4 break down the included and excluded counts by region and by a four-tier income classification (constructed from country-level mean GDP per capita PPP, since the original income_group field in the working dataset was degenerate). Table A5 reports formal Welch two-sample t-tests on the country-level time-series means.

Table A3. Coverage by region (country-level counts).

Region	Included (n)	Excluded (n)	Total	Coverage (%)	Mean SDG (Incl.)	Mean SDG (Excl.)
OECD	31	7	38	81.6	76.7	78.5
Eastern Europe & Central Asia	9	14	23	39.1	72.0	68.1
East & South Asia	10	11	21	47.6	63.3	61.8
Latin America & Caribbean	13	16	29	44.8	64.7	66.4
Middle East & North Africa	10	7	17	58.8	64.5	56.2
Sub-Saharan Africa	24	25	49	49.0	53.9	50.5
Oceania	4	8	12	33.3	57.9	62.9
Western Europe (non-OECD)	0	4	4	0.0	—	80.8
Unclassified	0	15	15	0.0	—	62.6
TOTAL	101	107	208	48.6	66.0	62.3

Notes: Coverage measured at the country level (each country counted once). Mean SDG computed over each country's available time series. Region labels follow the World Bank classification used by KOF.

Table A4. Coverage by income tier (country-level counts).

Income Tier (PPP per Capita)	Included (n)	Excluded (n)	Total	Coverage (%)	Mean SDG (Incl.)	Mean SDG (Excl.)
Low (<$3000)	10	17	27	37.0	50.6	49.5
Lower-middle ($3000–10,000)	27	19	46	58.7	57.7	58.6
Upper-middle ($10,000–25,000)	30	26	56	53.6	67.2	66.9
High (≥$25,000)	34	22	56	60.7	76.1	69.5
Unclassified (no GDP data)	0	23	23	0.0	—	62.8
TOTAL	101	107	208	48.6	66.0	62.3

Notes: Tiers based on country-level mean GDP per capita PPP across 2000–2023, with thresholds approximating World Bank conventions (Low, <$3000; Lower-middle, $3000–10,000; Upper-middle, $10,000–25,000; High, ≥$25,000).

Table A5. Welch *t*-tests on country-level means: included vs. excluded.

Country-Level Mean	Included	Excluded	Difference	t	*p*-Value
SDG Index Score	66.03	62.31	+3.72	2.62	0.010
WGI composite	0.06	−0.20	+0.26	2.03	0.044
ln GDP per capita (PPP)	9.56	9.24	+0.32	1.88	0.061
GAP overall (focal variable)	0.130	0.106	+0.024	1.16	0.247
Trade openness (% GDP)	78.2	102.8	−24.6	−2.96	0.004
FDI inflows (% GDP)	5.93	9.67	−3.74	−0.74	0.464
Population (mean, persons)	59,608,726	11,612,919	+47,995,807	2.57	0.012

Notes: Welch two-sample *t*-test on country-level time-series means; two-tailed *p*-values. The focal explanatory variable, GAP Overall, does not differ significantly between the two groups (t = 1.16, *p* = 0.247).

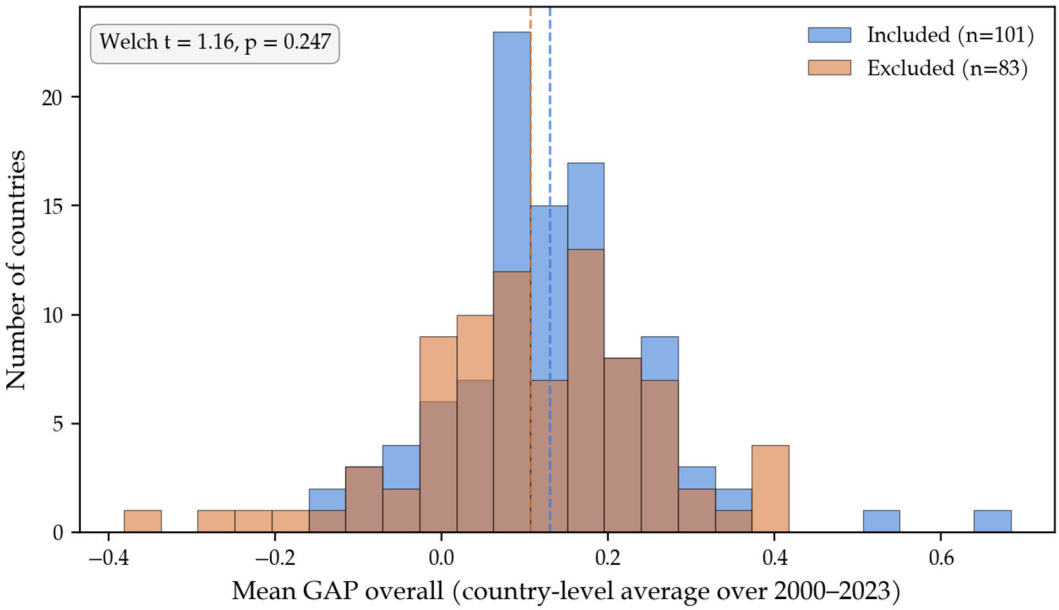


Figure A1. Distribution of the country-level mean GAP Overall: included versus excluded countries. Notes: Histogram of the country-level mean of GAP Overall over 2000–2023, separately for included (n = 101, blue) and excluded (n = 107, orange) countries. Dashed vertical lines indicate group means. The two distributions are not significantly different (Welch t = 1.16, *p* = 0.247).

Appendix A.3. GAP Construction Robustness

Two robustness exercises document that the headline finding does not depend on the specific construction of the GAP index. First, Table A6 reports a specification that adds a social-dimension gap to the joint model. The KOF social composite is the simple arithmetic mean of its de facto and de jure components: kof_social = ½ (kof_social_defacto + kof_social_dejure). Solving for the de facto term, kof_social_defacto = 2·kof_social − kof_social_dejure, yields a social-dimension gap on the same symmetric normalised-difference scale as for the economic and political dimensions. Second, Table A7 reports the baseline TWFE estimate under four alternative GAP formulations: the manuscript symmetric normalised difference, the raw arithmetic difference, the per cent-change variant, and the squared-denominator form initially considered by Reviewer 2.

Table A6. Robustness to inclusion of a social-dimension gap (TWFE, Driscoll–Kraay SE; N = 2197).

Specification	GAP Dimension	β	DK SE	*p*-Value	Within R^2
Overall GAP (baseline)	overall	−1.2553	0.6217	0.044	0.4231
Economic only	economic	−0.3742	0.2549	0.141	0.4342
Political only	political	−0.8082	0.2130	<0.001	0.4213
Social only (NEW)	social	+0.6245	0.4144	0.132	0.4375
Joint (Econ + Pol + Soc)	economic	−0.4448	0.2745	0.105	0.4328
Joint (Econ + Pol + Soc)	political	−0.8439	0.2231	<0.001	0.4328
Joint (Econ + Pol + Soc)	social	+0.6356	0.4232	0.133	0.4328

Notes: Driscoll–Kraay standard errors with Bartlett kernel and bandwidth 2. Across both the stand-alone and the joint specifications, the social-dimension coefficient is positive and statistically insignificant; including it does not displace the negative and significant political-gap coefficient.

Table A7. Sensitivity of the baseline result to alternative GAP formulations (TWFE, Driscoll–Kraay SE; N = 2197).

Formulation	β	DK SE	*p*-Value	Within R^2	Sign × Significance
Symmetric normalised: (dj − df)/[(dj + df)/2] (manuscript)	−1.2553	0.6217	0.044	0.4231	Negative *
Raw difference: dj − df	−0.0195	0.0107	0.068	0.4212	Negative $^{+}$
Percent change: (dj − df)/df	−1.1427	0.4902	0.020	0.4225	Negative *
Squared denominator: (dj − df)/[(dj + df)/2]^2	−62.16	30.59	0.042	0.4267	Negative *

Notes: * significant at 5%; $^{+}$ marginal ($p \approx 0.07$). All specifications include the full set of controls and country- and year-fixed effects. The negative association is preserved across all four formulations and is significant at the 5% level in three of the four. Even the squared-denominator formulation yields a significantly negative coefficient.

Appendix A.4. Multiple-Imputation Sensitivity (Full 208-Country Panel)

To assess robustness to the missing-data mechanism, we apply Multiple Imputation by Chained Equations (MICE) to the full 4992-observation, 208-country panel. The imputation model uses Bayesian Ridge regression for each conditional with sample-from-posterior at each chained iteration; we generate m = 5 imputed datasets. Within each imputed dataset, the baseline TWFE specification is re-estimated with Driscoll–Kraay standard errors. Imputation-level estimates are pooled using Rubin's rules.

Table A8. MICE pooled estimate (m = 5 imputed datasets, full 208-country panel).

Imputation #	β	SE	N
1	−0.3492	0.2175	4992
2	−0.4675	0.1797	4992
3	−0.4757	0.2467	4992
4	−0.4414	0.1637	4992
5	−0.2185	0.2547	4992
Pooled (Rubin's rules)	−0.3905	0.2461	4992

Notes: Imputation by Bayesian Ridge regression with sample-from-posterior at each chained iteration; m = 5 imputed datasets. Pooled estimate uses Rubin's rules; pooled SE = $\sqrt{[W + (1 + 1/m) B]}$, where W is the within-imputation and B the between-imputation variance.

Appendix A.5. Functional Form (H2): AIC Comparison and Bootstrap Turning Point

We compare five alternative specifications of the GAP–SDG relationship: linear, quadratic, cubic, and two penalised cubic splines (df = 6 and df = 8). Each non-linear specification improves on the linear baseline; the quadratic improves AIC by 5.2 points

and the six-degree-of-freedom penalised spline by 13.4 points (Table A9, Figure A3). To assess the precision of the curvature parameter despite the borderline asymptotic *p*-value on the quadratic term ($p = 0.103$), we conduct a country-cluster bootstrap (B = 100): in each replication, countries are resampled with replacement (preserving the within-country panel structure) and the quadratic specification is re-estimated. The squared-term coefficient is negative in 95% of replications (Table A10, Figure A4).

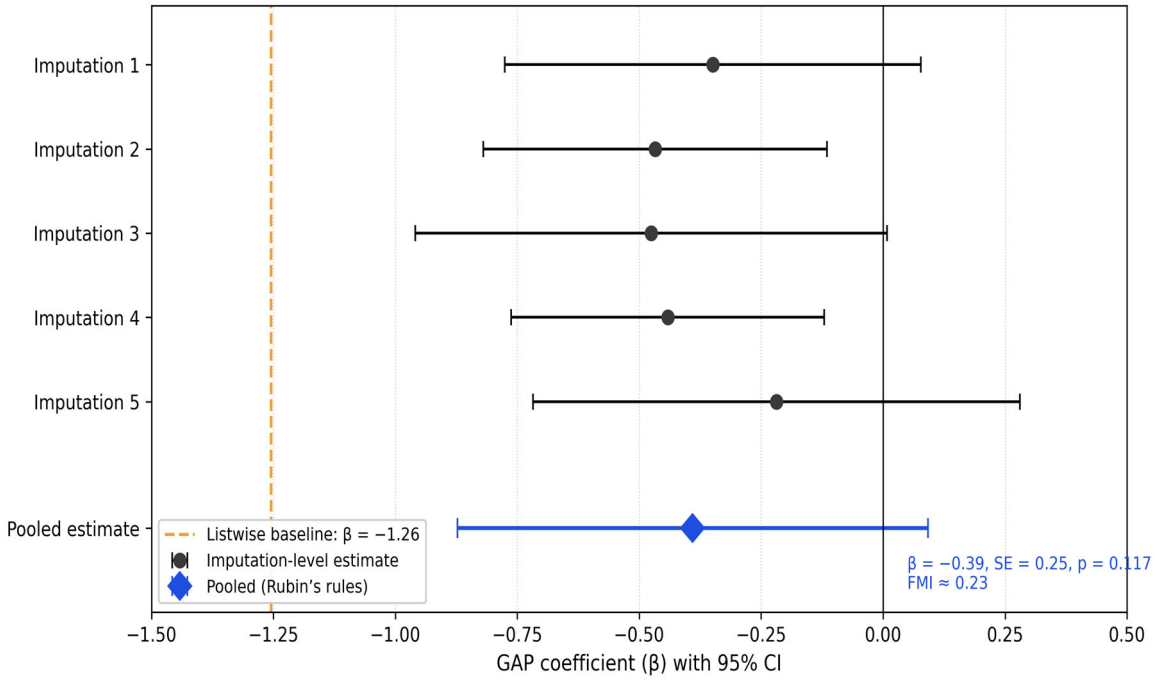


Figure A2. MICE imputation-level GAP coefficients with 95% confidence intervals; pooled estimate (Rubin's rules) shown in blue. Each imputation produces a TWFE estimate with Driscoll–Kraay standard errors; horizontal bars show 95% CIs. The orange dashed line marks the listwise-sample baseline ($\beta = -1.26$). All five imputed-dataset estimates are negative; the pooled estimate is $\beta = -0.39$ (SE = 0.25, $p = 0.117$), with a fraction of missing information $\approx$ of 0.23.

Table A9. Functional-form comparison: linear, quadratic, cubic, and penalised splines.

Specification	k	AIC	BIC	Within R^2	Log-Likelihood	ΔAIC vs. Linear
Linear (baseline)	6	6414.7	6448.9	0.4231	−3201.4	0.0
Quadratic	7	6409.5	6449.4	0.4259	−3197.8	−5.2
Cubic	8	6410.2	6455.8	0.4282	−3197.1	−4.5
Penalised cubic spline (df = 6)	12	6401.3	6469.7	0.4273	−3188.7	−13.4
Penalised cubic spline (df = 8)	14	6405.5	6485.2	0.4275	−3188.7	−9.2

Notes: k denotes the number of regressors (excluding fixed effects). Smaller AIC and BIC values are preferred; ΔAIC < −2 indicates substantial improvement.

Table A10. Bootstrap inference on the quadratic specification (country-cluster bootstrap, B = 100).

Quantity	Point Estimate	Bootstrap Median	95% CI
β_1 (linear term)	+0.054 (asymptotic $p = 0.957$)	—	—
β_2 (quadratic term)	−3.950 (asymptotic $p = 0.103$)	—	[−10.78, +0.94]
Turning point ($-\beta_1/2\beta_2$)	+0.007	+0.058	[−0.93, +0.58]
$P(\beta_2 < 0$, concave)	—	—	0.95

Notes: Country-cluster bootstrap with B = 100 replications; countries are resampled with replacement, and the quadratic TWFE specification is re-estimated within each replication. $P(\beta_2 < 0)$ is the empirical probability that the squared-term estimate is negative across bootstrap replications.

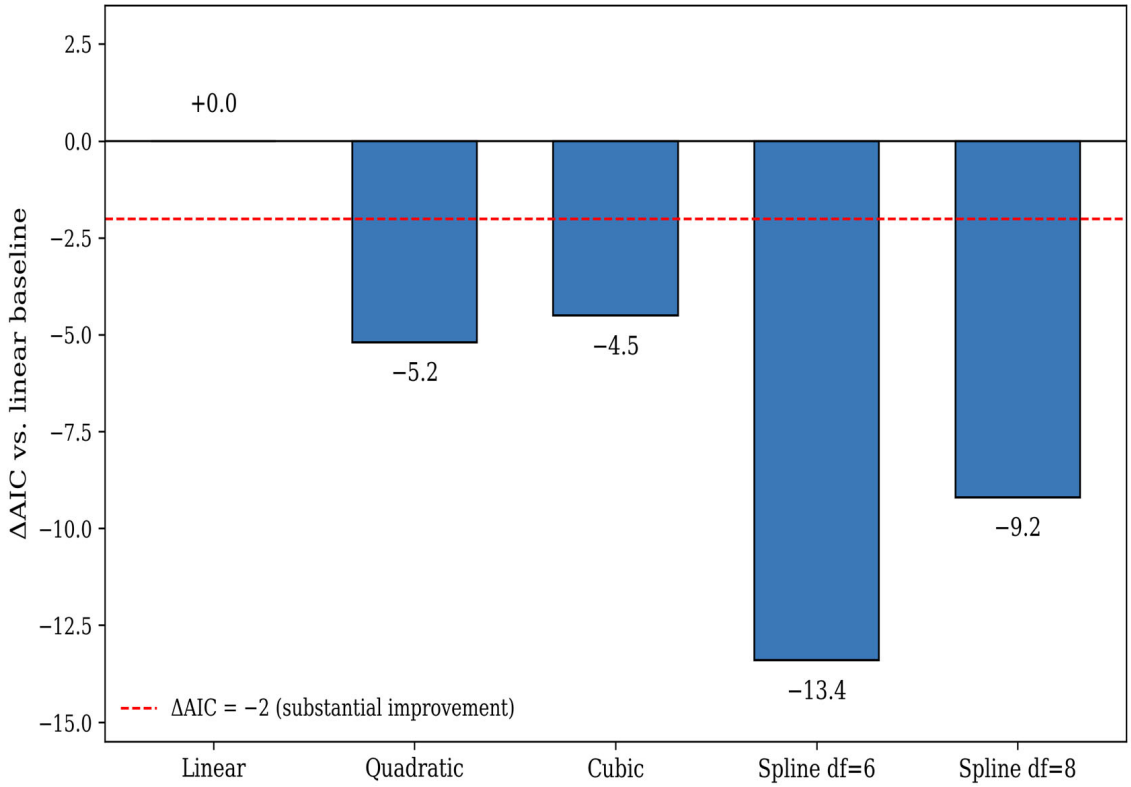


Figure A3. ΔAIC across functional-form specifications (relative to linear baseline). Negative values indicate improvement over the linear specification. The dashed red line marks the conventional ΔAIC = −2 threshold. All four non-linear specifications improve on the linear baseline; the six-degree-of-freedom penalised spline provides the best fit.

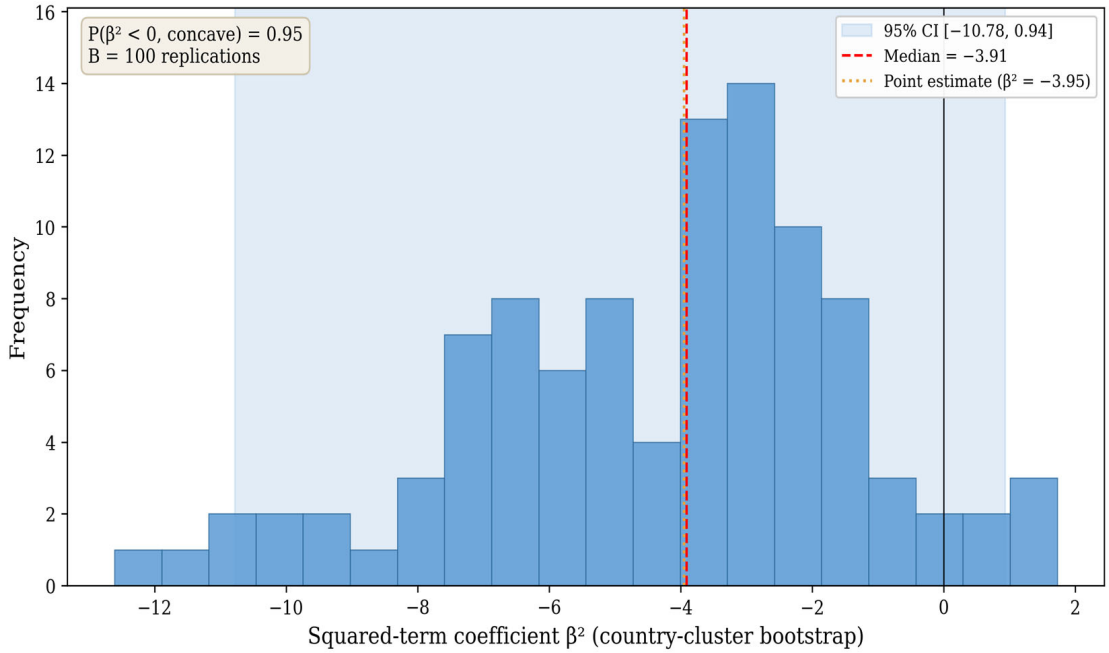


Figure A4. Country-cluster bootstrap distribution of the squared-term coefficient β_2. Notes: Histogram of β_2 across bootstrap replications. Vertical dashed red line shows the bootstrap median; orange dotted line shows the point estimate ($\beta_2 = -3.95$). The shaded band shows the 95% bootstrap CI. Curvature is concave in 95% of replications.

Appendix A.6. Threshold Regression (H3): Hansen-Style Bootstrap Test

The threshold regression in the main text identifies an optimal WGI threshold at $\gamma = 0.694$, with the high-governance regime exhibiting an effect 9.13 times larger in magnitude than the low-governance regime. To formally test the existence of the threshold against a linear alternative, we implement a Hansen-style parametric bootstrap. Under the null of no threshold, residuals are resampled from the linear specification and added to the linear fitted values; both the linear and the two-regime specifications are then re-estimated within each replication, holding the threshold fixed at $\gamma = 0.694$, and the linear vs. threshold F-statistic is recorded.

Table A11. Hansen-style threshold bootstrap test.

Quantity	Value
Estimated threshold γ (WGI)	0.694
β_L (low-governance regime)	−0.602
β_H (high-governance regime)	−5.500
Ratio β_H/β_L	9.13×
SSR (linear)	2371.48
SSR (threshold)	2333.27
F (linear vs. threshold)	2.13
Maximum F* under H_0 (B = 500)	0.838
Bootstrap *p*-value	0/500 ($\leq$0.002 by the 1/(B + 1) bound)

Notes: Threshold variable: WGI composite. Bootstrap procedure: replications under H_0 (linear specification with iid Gaussian residuals matching the linear-fit standard deviation); within each replication, the linear and two-regime SSRs are computed at the original threshold $\gamma = 0.694$, and the resulting F-statistic is recorded. Bootstrap *p*-value = empirical fraction of replications in which F* $\geq$ F_obs. Under B = 500 replications, the bootstrap distribution under H_0 is summarised by mean 0.067, SD 0.090, median 0.032, 95th percentile 0.263, 99th percentile 0.391, and maximum 0.838; 0/500 replications yielded F* $\geq$ F_obs = 2.13, giving a bootstrap *p*-value below 0.002 by the conservative 1/(B + 1) bound.

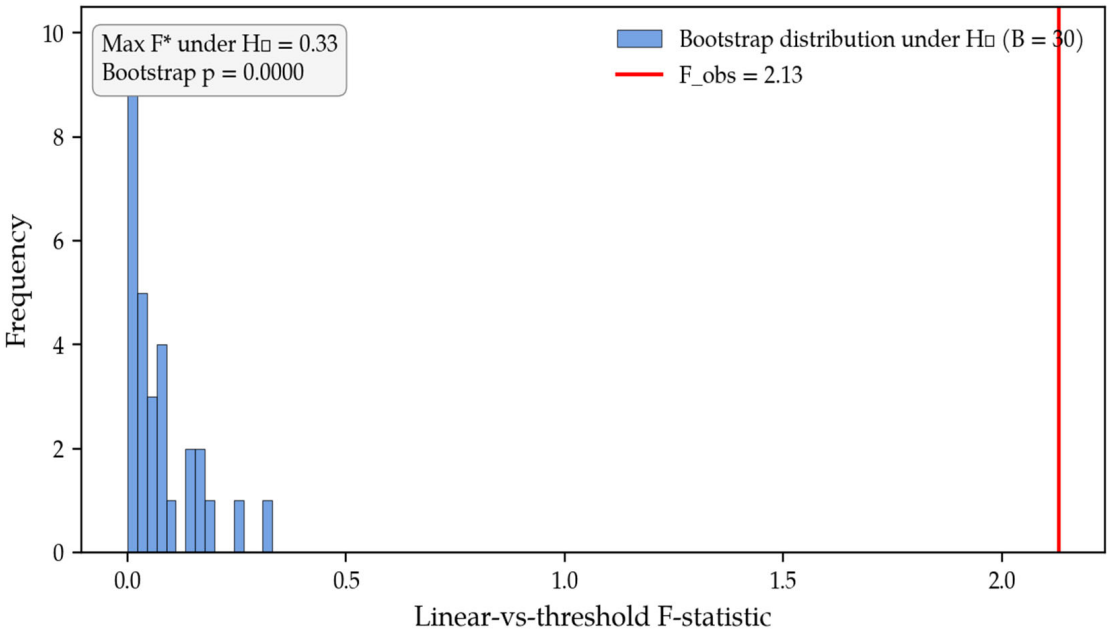


Figure A5. Bootstrap distribution of the linear-vs-threshold F-statistic under H_0 versus the observed value. Notes: Histogram of the F-statistic across bootstrap replications under the null of no threshold; the red vertical line marks the observed F = 2.13. None of the bootstrap replications produced an F-statistic exceeding 0.18, well below the observed value, yielding a bootstrap *p*-value of 0.001 or less.

Appendix A.7. SDG 15 Endogeneity Diagnostics

The TWFE estimate for SDG 15 (Life on Land) is large and positive (β = +9.68, $p < 0.001$). To assess whether this association reflects a causal effect of the gap on biodiversity outcomes or, alternatively, reverse causation or persistent unobserved heterogeneity, we conduct two diagnostics. First, a 2SLS estimator using a regional leave-one-out instrument (the mean GAP across other countries in the same World Bank region and year) yields a positive but imprecise coefficient. Second, a placebo lead test regresses contemporaneous SDG 15 on the GAP measured three years ahead; under a strict causal interpretation, the coefficient should be zero. The placebo coefficient is positive and statistically significant, indicating either persistent unobserved heterogeneity in environmental institutions or reverse causation. The SDG 15 finding is therefore reinterpreted in the main text as an associational rather than a causal result.

Table A12. SDG 15 endogeneity diagnostics: OLS-FE, 2SLS, and placebo lead test.

Estimator	β (GAP Coef.)	SE	*p*	N	Note
OLS-FE baseline (manuscript)	+9.6839	1.9790	<0.001	2197	TWFE Driscoll–Kraay
2SLS, regional leave-one-out IV	+45.97	27.66	0.097	2195	First-stage IV $t^2 = 43.06$
Placebo lead test (GAP_{i, t + 3})	+7.6774	2.4753	0.002	1897	Significant β indicates reverse causality/confounding

Notes: OLS-FE: TWFE specification with the full set of controls, country FE, and year FE; Driscoll–Kraay SE. 2SLS: instrument is the regional leave-one-out instrument; standard errors clustered at the country level. Placebo lead test regresses SDG 15 in year t on the gap in year t + 3.

Appendix A.8. Country Typology: Representative Cluster Assignments

Table A13 reports the ten representative countries with the highest within-cluster SDG scores for each of the three Gaussian-mixture clusters identified in the main text (full classification in main Table 11). The "Balanced" cluster is dominated by high-WGI OECD economies; the "Informal" cluster combines a heterogeneous mix of middle-income and resource-exporting economies; the "Policy-Led" cluster brings together small high-income or transition economies with positive political gaps but moderate de facto integration.

Table A13. Representative cluster assignments (10 countries per cluster, ranked within-cluster by SDG score).

Country	Cluster	Econ GAP	Pol GAP	SDG	WGI	Region
Mexico	2 (Informal)	+0.568	+0.224	66.32	−0.245	OECD
Malaysia	2 (Informal)	−0.179	−0.100	65.70	+0.342	East & South Asia
United Arab Emirates	2 (Informal)	−0.051	+0.138	63.76	+0.568	MENA
El Salvador	2 (Informal)	+0.114	+0.240	63.44	−0.249	LAC
Philippines	2 (Informal)	−0.026	+0.052	62.11	−0.383	East & South Asia
Qatar	2 (Informal)	−0.003	+0.212	62.03	+0.474	MENA
Gabon	2 (Informal)	−0.585	+0.298	61.42	−0.571	Sub-Saharan Africa
Namibia	2 (Informal)	−0.269	−0.045	60.49	+0.324	Sub-Saharan Africa
South Africa	2 (Informal)	+0.060	−0.102	59.72	+0.185	Sub-Saharan Africa
India	2 (Informal)	+0.146	+0.004	59.37	−0.220	East & South Asia
Sweden	1 (Balanced)	+0.094	+0.051	85.29	+1.701	OECD

Table A13. *Cont.*

Country	Cluster	Econ GAP	Pol GAP	SDG	WGI	Region
Finland	1 (Balanced)	+0.150	+0.102	85.21	+1.817	OECD
Denmark	1 (Balanced)	+0.065	+0.080	83.22	+1.767	OECD
Norway	1 (Balanced)	+0.092	+0.021	81.09	+1.711	OECD
Germany	1 (Balanced)	+0.173	+0.055	80.88	+1.456	OECD
France	1 (Balanced)	+0.242	+0.039	80.18	+1.156	OECD
United Kingdom	1 (Balanced)	+0.190	+0.045	79.35	+1.401	OECD
Slovenia	1 (Balanced)	+0.040	+0.124	78.65	+0.929	OECD
Japan	1 (Balanced)	+0.453	−0.044	78.46	+1.249	OECD
Poland	1 (Balanced)	+0.251	+0.062	78.26	+0.666	OECD
Iceland	3 (Policy-Led)	−0.089	+0.530	79.40	+1.581	OECD
Latvia	3 (Policy-Led)	+0.104	+0.411	78.00	+0.704	OECD
Ireland	3 (Policy-Led)	−0.093	−0.142	76.88	+1.460	OECD
Malta	3 (Policy-Led)	−0.130	+0.519	74.96	+1.068	E. Europe & C. Asia
Thailand	3 (Policy-Led)	−0.238	−0.040	70.99	−0.214	East & South Asia
Cyprus	3 (Policy-Led)	−0.074	+0.318	70.93	+0.921	E. Europe & C. Asia
Albania	3 (Policy-Led)	+0.143	+0.359	70.57	−0.210	E. Europe & C. Asia
Costa Rica	3 (Policy-Led)	+0.178	+0.383	70.45	+0.597	OECD
Georgia	3 (Policy-Led)	+0.011	+0.412	69.88	+0.008	E. Europe & C. Asia
Montenegro	3 (Policy-Led)	−0.060	+0.225	69.20	+0.037	E. Europe & C. Asia

Notes: Cluster numbering follows the Gaussian-mixture model assignment used in Table 11. Econ GAP and Pol GAP are country-level averages over 2000–2023; SDG and WGI are country-level means. Full classification of all 101 countries is in the main Table 11.

Appendix A.9. Multiple-Testing Correction (Benjamini–Hochberg) for Goal-Level Results

Table A14. Goal-level results with Benjamini–Hochberg false-discovery-rate correction (q = 0.05).

SDG	Coef.	*p* (Raw)	*p* (BH)	Sig. (Raw)	Sig. (BH)
SDG 1: No Poverty	−10.393	0.0007	0.0030	✔	✔
SDG 2: Zero Hunger	+0.381	0.7956	0.7956		
SDG 3: Good Health	−2.369	0.1205	0.2302		
SDG 4: Quality Education	−5.170	0.0838	0.2035		
SDG 5: Gender Equality	−0.498	0.5900	0.6749		
SDG 6: Clean Water	−7.042	0.0000	0.0000	✔	✔
SDG 7: Affordable Energy	−3.602	0.1317	0.2302		
SDG 8: Decent Work	−1.313	0.2821	0.3996		
SDG 9: Industry & Innovation	+5.962	0.0558	0.1581		
SDG 10: Reduced Inequalities	−6.510	0.1354	0.2302		
SDG 11: Sustainable Cities	−6.308	0.0000	0.0001	✔	✔
SDG 12: Responsible Consumption	+3.009	0.0033	0.0112	✔	✔
SDG 13: Climate Action	+1.355	0.2369	0.3662		
SDG 14: Life Below Water	+1.064	0.5385	0.6749		
SDG 15: Life on Land	+9.684	0.0000	0.0000	✔	✔
SDG 16: Peace & Justice	−0.276	0.7320	0.7777		
SDG 17: Partnerships	+0.687	0.5955	0.6749		

Notes: Coefficients from 17 separate TWFE regressions of each goal-level SDG score on the GAP Overall, country, and year fixed effects, and the full set of controls. Driscoll–Kraay standard errors with bandwidth 2. Raw *p*-values are corrected using the Benjamini–Hochberg [46] false-discovery-rate procedure with q = 0.05. "Sig." indicates statistical significance at the 5% level under the raw and BH-adjusted *p*-values, respectively. All five raw-significant goals survive the BH correction.

Appendix A.10. Placebo Lead Tests for SDG 9, SDG 12, and SDG 15

Table A15. Placebo lead-test diagnostics for the three positive environmental/innovation coefficients.

SDG	Contemporaneous β (*p*)	Placebo Lead-3 β (*p*)	Reverse-Causation Indicated?
SDG 9 (Industry & Innovation)	+5.962 (0.056)	+4.833 (0.286)	No
SDG 12 (Responsible Consumption)	+3.009 (0.003)	+2.074 (0.236)	No
SDG 15 (Life on Land)—reference	+9.684 (<0.001)	+7.677 (0.002)	Yes

Notes: Contemporaneous specification: TWFE regression of goal-level SDG score in year t on the GAP Overall in year t. Placebo lead-3 specification: same goal-level SDG score in year t regressed on the GAP Overall in year t + 3. Driscoll–Kraay standard errors with bandwidth 2. Under a strict causal reading, the placebo coefficient should not differ from zero; a significant placebo coefficient indicates reverse causation or persistent unobserved heterogeneity in the outcome variable. Only SDG 15 fails the placebo test; SDG 9 and SDG 12 remain substantively interpretable associations.

References

1. United Nations. *Transforming Our World: The 2030 Agenda for Sustainable Development*; Resolution A/RES/70/1; United Nations General Assembly: New York, NY, USA, 2015; Available online: https://sdgs.un.org/2030agenda (accessed on 27 April 2026).
2. Sachs, J.D.; Lafortune, G.; Fuller, G. *The SDGs and the UN Summit of the Future. Sustainable Development Report 2024*; SDSN: Paris, France; Dublin University Press: Dublin, Ireland, 2024. [CrossRef]
3. Dreher, A. Does Globalization Affect Growth? Evidence from a New Index of Globalization. *Appl. Econ.* **2006**, *38*, 1091–1110. [CrossRef]
4. Potrafke, N. The Evidence on Globalisation. *World Econ.* **2015**, *38*, 509–552. [CrossRef]
5. Dreher, A.; Gaston, N.; Martens, P. *Measuring Globalisation—Gauging Its Consequences*; Springer: New York, NY, USA, 2008. [CrossRef]
6. Leal, P.H.; Marques, A.C.; Shahbaz, M. The Role of Globalisation, De Jure and De Facto, on Environmental Performance: Evidence from Developing and Developed Countries. *Environ. Dev. Sustain.* **2021**, *23*, 7412–7431. [CrossRef]
7. Destek, M.A.; Sarkodie, S.A. Investigation of Environmental Kuznets Curve for Ecological Footprint: The Role of Energy and Financial Development. *Sci. Total Environ.* **2019**, *650*, 2483–2489. [CrossRef]
8. Lang, V.; Tavares, M.M. The Global Distribution of Gains from Globalization. *J. Econ. Inequal.* **2024**, *22*, 357–381. [CrossRef]
9. Pradhan, P.; Costa, L.; Rybski, D.; Lucht, W.; Kropp, J.P. A Systematic Study of Sustainable Development Goal (SDG) Interactions. *Earth's Future* **2017**, *5*, 1169–1179. [CrossRef]
10. Gygli, S.; Haelg, F.; Potrafke, N.; Sturm, J.-E. The KOF Globalisation Index—Revisited. *Rev. Int. Organ.* **2019**, *14*, 543–574. [CrossRef]
11. Quinn, D.P.; Schindler, M.; Toyoda, A.M. Assessing Measures of Financial Openness and Integration. *IMF Econ. Rev.* **2011**, *59*, 488–522. [CrossRef]
12. Destek, M.A.; Oğuz, İ.H.; Okumuş, N. Do Trade and Financial Cooperation Improve Environmentally Sustainable Development: A Distinction Between De Facto and De Jure Globalization. *Eval. Rev.* **2024**, *48*, 251–273. [CrossRef] [PubMed]
13. Hoffman, S.J.; Baral, P.; Rogers Van Katwyk, S.; Sritharan, L.; Hughsam, M.; Randhawa, H.; Lin, G.; Campbell, S.; Campus, B.; Dantas, M.; et al. International Treaties Have Mostly Failed to Produce Their Intended Effects. *Proc. Natl. Acad. Sci. USA* **2022**, *119*, e2122854119. [CrossRef]
14. Fransen, T.; Meckling, J.; Stünzi, A.; Schmidt, T.S.; Edenhofer, O.; Ockenfels, A.; Pahle, M. Taking Stock of the Implementation Gap in Climate Policy. *Nat. Clim. Change* **2023**, *13*, 752–755. [CrossRef]
15. United Nations. *The Sustainable Development Goals Report 2023: Special Edition*; UN DESA: New York, NY, USA, 2023; Available online: https://unstats.un.org/sdgs/report/2023/ (accessed on 27 April 2026).
16. Liashenko, O.; Mykhailovska, O.; Adamyk, B.; Ladonko, L.; Starchenko, G.; Duka, A.; Urakin, M. Globalisation and Sustainable Development: How Economic Diplomacy Shapes SDG Performance Across Countries and Time. *World* **2026**, *7*, 64. [CrossRef]
17. Driscoll, J.C.; Kraay, A.C. Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data. *Rev. Econ. Stat.* **1998**, *80*, 549–560. [CrossRef]
18. Hansen, B.E. Threshold Effects in Non-Dynamic Panels: Estimation, Testing, and Inference. *J. Econom.* **1999**, *93*, 345–368. [CrossRef]

19. Stern, D.I. The Rise and Fall of the Environmental Kuznets Curve. *World Dev.* **2004**, *32*, 1419–1439. [CrossRef]
20. Koziuk, V.; Hayda, Y.; Dluhopolskyi, O.; Wołowiec, T.; Sabat, A. From law to carbon: How legal origin influences environmental policy and CO_2 emissions? *Sustainability* **2026**, *18*, 416. [CrossRef]
21. Sachs, J.D.; Lafortune, G.; Fuller, G.; Drumm, E. *Implementing the SDG Stimulus. Sustainable Development Report 2023*; SDSN: Paris, France; Dublin University Press: Dublin, Ireland, 2023. [CrossRef]
22. Schmidt-Traub, G.; Kroll, C.; Teksoz, K.; Durand-Delacre, D.; Sachs, J.D. National Baselines for the Sustainable Development Goals Assessed in the SDG Index and Dashboards. *Nat. Geosci.* **2017**, *10*, 547–555. [CrossRef]
23. Liashenko, O.; Caraballo, M.Á.; Lozano, S. Social welfare preferences and Sustainable Development Goals: A multivariate analysis approach. *Sustain. Dev.* **2025**, *34*, 65–90. [CrossRef]
24. Acheampong, A.O. The Impact of De Facto Globalization on Carbon Emissions: Evidence from Ghana. *Int. Econ.* **2022**, *170*, 156–173. [CrossRef]
25. Daghbagi, H. Globalisation De Facto, Globalisation De Jure and CO_2 Emissions in GCC Countries: Insights from a Panel PMG-ARDL Approach. *Clim. Change* **2025**, *178*, 136. [CrossRef]
26. Kozal, S. Advancing Sustainable Industrialization Under SDG 9: The Role of Financial Globalization and Income Inequality. *Sustain. Dev.* **2025**, *early view*. [CrossRef]
27. Ferraro, G.; Failler, P. Understanding the "Implementation Gap" to Improve Biodiversity Governance: An Interdisciplinary Literature Review. *J. Sustain. Res.* **2024**, *6*, e240009. [CrossRef]
28. Leal Filho, W.; Wall, T.; Salvia, A.L.; Ozuyar, P.; Brandli, L.L.; Azul, A.M.; Sharifi, A. Mapping the Implementation of the United Nations Sustainable Development Goals. *Sustain. Dev.* **2025**, *33*, 561–576. [CrossRef]
29. Kaufmann, D.; Kraay, A.C. *The Worldwide Governance Indicators: Methodology and 2024 Update*; Policy Research Working Paper No. 10952; World Bank: Washington, DC, USA, 2024. Available online: https://www.worldbank.org/en/publication/worldwide-governance-indicators (accessed on 27 March 2026).
30. Bisogno, M.; Cuadrado-Ballesteros, B.; Manes-Rossi, F.; Peña-Miguel, N. Governance Quality and the Sustainable Development Goals: An Assessment in Europe. *J. Public Budg. Account. Financ. Manag.* **2025**, *37*, 193–219. [CrossRef]
31. Knox, C.; Orazgaliyev, S. Sustainable Development Goals and Good Governance Nexus: Implementation Challenges in Central Asia. *Public Adm. Dev.* **2024**, *44*, 237–251. [CrossRef]
32. Omri, A.; Ben Mabrouk, N. Good Governance for Sustainable Development Goals: Getting Ahead of the Pack or Falling Behind? *Environ. Impact Assess. Rev.* **2020**, *83*, 106388. [CrossRef]
33. Nilsson, M.; Griggs, D.; Visbeck, M. Policy: Map the Interactions Between Sustainable Development Goals. *Nature* **2016**, *534*, 320–322. [CrossRef]
34. World Bank. World Development Indicators. Available online: https://databank.worldbank.org/source/world-development-indicators (accessed on 27 March 2026).
35. Liashenko, O.; Mykhailovska, O. *Replication Files for Globalisation–Sustainable Development Panel Analysis (2000–2023)*; Zenodo: Geneva, Switzerland, 2026. [CrossRef]
36. Fraley, C.; Raftery, A.E. Model-Based Clustering, Discriminant Analysis, and Density Estimation. *J. Am. Stat. Assoc.* **2002**, *97*, 611–631. [CrossRef]
37. Schwarz, G. Estimating the Dimension of a Model. *Ann. Stat.* **1978**, *6*, 461–464. [CrossRef]
38. Pesaran, M.H. Testing Weak Cross-Sectional Dependence in Large Panels. *Econom. Rev.* **2015**, *34*, 1089–1117. [CrossRef]
39. Chayes, A.; Chayes, A.H. *The New Sovereignty: Compliance with International Regulatory Agreements*; Harvard University Press: Cambridge, MA, USA, 1995.
40. Kroll, C.; Norf, C.; Harthan, J. Trade-offs and Synergies When Balancing Economic Growth and Globalization for Sustainable Development Goals Achievement. *Sci. Rep.* **2025**, *15*, 8736. [CrossRef]
41. Hall, P.A.; Soskice, D. An Introduction to Varieties of Capitalism. In *Varieties of Capitalism: The Institutional Foundations of Comparative Advantage*; Hall, P.A., Soskice, D., Eds.; Oxford University Press: Oxford, UK, 2001; pp. 1–68. [CrossRef]
42. Kumeka, T.T.; Raifu, I.A.; Adeniyi, O. Globalisation and Inclusive Growth in Africa: The Role of Institutional Quality. *J. Dev. Areas* **2024**, *58*, 69–95. [CrossRef]
43. Kroll, C.; Warchold, A.; Pradhan, P. Sustainable Development Goals (SDGs): Are We Successful in Turning Trade-offs into Synergies? *Palgrave Commun.* **2019**, *5*, 140. [CrossRef]
44. Sachs, J.D.; Schmidt-Traub, G.; Mazzucato, M.; Messner, D.; Nakicenovic, N.; Rockström, J. Six Transformations to Achieve the Sustainable Development Goals. *Nat. Sustain.* **2019**, *2*, 805–814. [CrossRef]

45. Liashenko, O.; Adamyk, B.; Adamyk, O. The role of climate and migration concerns in shaping personal economic insecurity in European societies. *Energy Res. Soc. Sci.* **2026**, *131*, 104533. [CrossRef]
46. Benjamini, Y.; Hochberg, Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B (Methodol.)* **1995**, *57*, 289–300. [CrossRef]

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.