Measuring Support for the International Liberal Order: A Multidimensional Dynamic IRT Approach

[Author Names Redacted]

February 2026

Abstract

How has systemic support for the International Liberal Order (ILO) evolved since the end of the Cold War? Existing approaches rely on UN General Assembly voting or expert surveys, which capture geopolitical alignment but not the institutional commitments that constitute the order itself. We propose a direct measurement strategy that combines treaty ratification data across six domains — investment, security, environment, human rights, arms control, and intellectual property — with tariff data for a seventh domain, trade. Five treaty domains extend through 2024 (seven five-year periods from 1990 to 2024); the security domain is constrained to 2018 by the coverage of the ATOP alliance dataset. Trade data from the WITS/TRAINS database cover 176 countries from 1990 to 2022. For the six treaty domains, we estimate a two-dimensional dynamic item response theory model separately for each, recovering latent ideal points that distinguish ILO support from domain-specific confounders. We complement these per-domain models with a hierarchical joint IRT model estimated via Bayesian MCMC that decomposes country positions into a common ILO-support factor and domain-specific deviations across four domains, using V-Dem's liberal democracy index as a predictor. For trade, we use directly observed MFN tariff rates and preferential margins to distinguish three ideal types: universally open, club-open, and closed countries. Three findings emerge. First, the investment domain shows clear deepening of ILO commitments that continues into the 2020–2024 period, while security alliances drift away from the ILO pole. Second, trade tariffs fall dramatically across three decades, but preferential margins widen, indicating a shift from multilateral to club-based liberalization. Third, the standard deviation of ideal points rises across domains, suggesting that the post-Cold War order is characterized not by collapse or convergence but by growing heterogeneity in countries' institutional commitments. The hierarchical model confirms that domestic liberal democracy predicts aggregate ILO support and reveals that investment treaty engagement is largely idiosyncratic, while human rights and arms control are more tightly linked to the common factor.

Introduction

The state of the International Liberal Order (ILO) has become one of the central questions in international relations. A growing literature debates whether the post-Cold War order is in decline (Maull 2018; Cooley and Nexon 2020), undergoing contestation from within and without (Börzel and Zürn 2021), or being reimagined by non-Western actors with alternative visions of legitimate governance (Loh, Paes, and Zarakol 2025). Scholars have documented how populist governments challenge multilateral institutions through voting behavior (Destradi and Vüllers 2024) and how the liberal order's own institutional architecture can become self-undermining (Farrell and Newman 2021).

Yet despite the richness of this theoretical debate, the empirical measurement of ILO support remains underdeveloped. The most common quantitative approach uses UN General Assembly voting patterns to estimate country positions on a liberal-versus-illiberal spectrum (Bailey, Strezhnev, and Voeten 2017; Voeten 2021). While informative about geopolitical alignment, UNGA votes capture cheap talk — non-binding resolutions that impose no material costs. An alternative approach uses expert surveys such as V-Dem's liberal democracy indices, which measure domestic institutional quality rather than international commitments. Still others have developed dynamic measures of specific domains, such as Fariss's (2014) latent human rights protection scores, which account for changing standards of accountability.

We argue that the ILO is constituted, in significant part, by the web of international treaties that countries ratify and the trade policies they adopt: bilateral investment treaties, security alliances, environmental agreements, human rights conventions, arms control accords, intellectual property instruments, and tariff regimes. Treaty ratification entails legal obligations, domestic policy adjustments, and reputational stakes. Tariff policy directly reflects a country's openness to the multilateral trading system. A country's pattern of treaty ratification and trade liberalization across these seven domains therefore provides a measure of its engagement with the institutional architecture of the liberal order. We acknowledge that ratification is an imperfect proxy for commitment: a substantial literature documents the gap between ratification and compliance (Hathaway 2002; Simmons 2009), and recent work emphasizes the role of monitoring and enforcement mechanisms in determining whether treaty commitments translate into behavioral change (Kelley and Simmons 2015). Our measure should therefore be understood as capturing formal institutional engagement — the willingness to accept legal obligations — rather than revealed preferences over compliance with liberal norms.

This paper makes four contributions. First, we construct a comprehensive dataset of treaty ratification events across six domains for 206 countries over seven five-year periods (1990–2024) for five domains and six periods (1990–2018) for security, complemented by tariff data for a seventh domain (trade) covering 176 countries through 2022. Second, we estimate a two-dimensional dynamic item response theory (IRT) model for each treaty domain, recovering latent ideal points that separate ILO support from domain-specific confounders. For trade, we use directly observed tariff measures — MFN rates and preferential margins — that distinguish universally open, club-open, and closed countries. Third, we complement the per-domain models with a hierarchical joint IRT model that decomposes country positions into a common ILO-support factor and domain-specific deviations, incorporating V-Dem's liberal democracy index as a predictor and domain-specific covariates (tax haven status, nuclear weapon state status) to explain cross-domain variation. Fourth, we characterize the aggregate evolution of ILO support, showing that the post-Cold War period is marked not by uniform decline but by increasing differentiation.

Data

We assemble data from seven domains of international cooperation, each representing a distinct pillar of the ILO. Six domains are measured through treaty ratification; the seventh (trade) is measured directly through tariff data.

Treaty Sources

Table 1 lists the six treaty-based domains and their primary sources.

Table 1. Treaty domains and data sources

Domain	Source	Treaties	Coverage
Investment	UNCTAD IIA Navigator	2,874 BITs	1959–2024
Security	ATOP 5.1	874 alliances	1815–2018
Environment	IEADB	3,111 IEAs	1857–2024
Human Rights	UN Treaty Collection Ch. IV	18 conventions	1965–2024
Arms Control	UN Treaty Collection Ch. XXVI + NPT/BWC	11 treaties	1968–2024
Intellectual Property	WIPO Lex	26 treaties	1883–2024

Notes: Treaty counts refer to treaty-period items in the flow matrix (after period expansion). Investment treaties are predominantly bilateral investment treaties. Security alliances are from the Alliance Treaty Obligations and Provisions dataset (Leeds et al. 2002), which ends in 2018; the remaining five domains extend through 2024. Environmental agreements are from the International Environmental Agreements Database (Mitchell et al. 2020). Arms control includes the NPT, BWC, CWC, CTBT, Ottawa Convention, and six additional UN Chapter XXVI instruments, plus the Treaty on the Prohibition of Nuclear Weapons (TPNW). The human rights domain includes 9 core UN conventions and their optional protocols. The intellectual property domain includes 26 WIPO-administered treaties.

The six treaty domains differ markedly in their structure. Investment and security are dominated by bilateral agreements, producing thousands of treaty-period items but sparse ratification matrices (most country-treaty pairs are zeros under flow coding). Environment spans both bilateral and large multilateral agreements, yielding the richest dataset. Human rights and arms control consist entirely of large multilateral conventions with near-universal membership, providing the starkest contrasts between early adopters and holdouts.

Trade Data

Treaty-based measurement is unavailable for trade because the key trade agreements (GATT/WTO, regional FTAs) are too few and heterogeneous to support an IRT model. Instead, we measure trade engagement directly through tariff data from the World Integrated Trade Solution (WITS) database, which draws on UNCTAD's TRAINS (Trade Analysis Information System) dataset.

For each of 176 countries, we collect applied MFN (Most Favored Nation) tariff rates and bilateral preferential tariff rates at the HS 4-digit level for all available years from 1990 to 2022. We then compute two summary measures per country-year:

1. **Mean MFN tariff**: the simple average across 21 HS sections (Harmonized System chapters grouped by WCO classification), giving equal weight to each section.
2. **Mean preferential margin**: for each HS4 product, the margin is defined as the MFN tariff minus the lowest preferential tariff available to any partner; we take the maximum margin across partners for each product, then average across sections.

These two measures capture distinct aspects of trade engagement with the ILO. The MFN tariff reflects a country's general openness to the multilateral trading system — low MFN tariffs indicate broad liberalization accessible to all trading partners. The preferential margin captures the depth of preferential agreements — a high margin indicates that the country offers substantially lower tariffs to select partners relative to the MFN baseline. Together, the two measures distinguish three ideal types: *universally open* countries (low MFN, low margin), *club-open* countries (high MFN, high margin), and *closed* countries (high MFN, low margin).

Flow Coding

For each domain, we construct a binary response matrix $\mathbf{Y}$ of dimension $N \times J$, where N is the number of countries and J is the number of treaty-period items. We employ **flow coding**: a country receives a $+1$ for a treaty-period item if it newly ratified that treaty during the period, and -1 if the treaty was available but the country did not ratify it. Missing values (treaty not yet open, or country not yet independent) are coded as 0 and excluded from the likelihood.

Flow coding captures *new commitments* in each period, providing a dynamic signal of engagement. We use seven five-year periods (1990–1994 through 2020–2024) for five domains and six periods (1990–1994 through 2015–2018) for security, which is constrained by the ATOP dataset's coverage. As a robustness check, we also estimate models under **stock coding**, where a country is coded $+1$ in all periods during which it is a party to the treaty (Section 5).

Table 2 summarizes the estimation data by domain.

Table 2. Data summary by issue domain

Domain	N	J	T	Bilateral	Mini (4–10)	Medium (11–50)	Large (51+)
Investment	204	2,874	7	85.3%	2.5%	2.7%	0.0%
Security	164	935	6	89.1%	3.0%	2.5%	0.5%
Environmen	206	3,111	7	6.1%	23.8%	24.7%	32.9%
Human Rights	197	98	7	0.0%	0.0%	5.6%	94.4%
Arms Control	195	56	7	0.0%	0.0%	0.0%	100.0%
Intellectual Property	186	152	7	0.0%	3.8%	23.1%	73.1%

Notes: N = number of countries; J = number of treaty-period items in the flow matrix; T = number of five-year periods (7 for 1990–2024, 6 for security which ends in 2018). Size columns report the share of base treaties by all-time membership: Bilateral = 2 parties; Mini = 4–10; Medium = 11–50; Large = 51+.

Model

For the six treaty-based domains, we estimate a two-dimensional dynamic IRT model that recovers latent ideal points from binary ratification data. The trade domain, measured through continuous tariff data, is analyzed descriptively using the MFN and preferential margin measures defined above

rather than through a latent variable model. This section describes the IRT specification applied to the treaty domains.

Two-Dimensional Dynamic IRT

We estimate a two-dimensional dynamic probit IRT model separately for each treaty domain. Let $y_{ij} \in \{+1, -1\}$ denote the observed ratification decision of country i on treaty-period item j, and let $s(j) \in \{1, \ldots, T\}$ denote the period associated with item j. The observation model is:

$$P(y_{ij} = +1 \mid \mathbf{x}_{i,s(j)}) = \Phi(\alpha_j + \beta_j' \mathbf{x}_{i,s(j)})$$

where Φ is the standard normal CDF, α_j is the item intercept (difficulty), $\beta_j = (\beta_{j1}, \beta_{j2})'$ is the two-dimensional discrimination vector, and $\mathbf{x}_{it} = (x_{it1}, x_{it2})'$ is the ideal point of country i in period t.

Ideal points evolve as a random walk:

$$\mathbf{x}_{i,1} \sim N(\mu_{0i}, \Sigma_{0i}), \qquad \mathbf{x}_{i,t} \mid \mathbf{x}_{i,t-1} \sim N(\mathbf{x}_{i,t-1}, \Omega)$$

where $\Omega = \omega^2 \mathbf{I}_2$ governs the smoothness of temporal evolution. We set $\omega^2 = 0.1$ as the baseline, following the convention in the emIRT package (Imai, Lo, and Olmsted 2016), and vary it across a 50-fold range in robustness checks.

Identification

A K-dimensional probit IRT model has $K(K+1)$ degrees of rotational and translational invariance. For $K = 2$, this requires 6 constraints. We achieve identification by fixing the positions of $K+1 = 3$ anchor countries per domain using tight Gaussian priors ($\sigma = 0.01$). Each anchor provides $K = 2$ constraints, yielding $3 \times 2 = 6$ total constraints.

Anchor countries are selected based on two criteria: (1) the first dimension should be oriented so that countries with extensive treaty ratification records — which we interpret as reflecting greater ILO support — load positively, and countries with minimal ratification load negatively, and (2) the third anchor should load on the second dimension, capturing domain-specific variation orthogonal to the first dimension. We emphasize that the substantive label "ILO support" for dim-1 is imposed through anchor placement rather than recovered endogenously from the data; the latent dimension is properly understood as the axis defined by the chosen anchor countries, and the ILO-support interpretation depends on the validity of those anchor choices (see Robustness, Section R2, for anchor sensitivity analysis). Table 3 reports the anchor configurations.

Table 3. Anchor countries by domain

Domain	dim-1 (+)	dim-1 (−)	dim-2
Investment	DNK (+2, 0)	IRN (−2, 0)	CHN (0, −2)
Security	DNK (+2, 0)	IRN (−2, 0)	UKR (0, −2)
Environment	DNK (+2, 0)	SAU (−2, 0)	AUS (0, −2)
Human Rights	DNK (+2, 0)	PRK (−2, 0)	USA (0, −2)
Arms Control	NZL (+2, 0)	ISR (−2, 0)	IND (0, −2)

Domain	dim-1 (+)	dim-1 (−)	dim-2
Intellectual Property	DNK (+2, 0)	AGO (−2, 0)	BRA (0, −2)

Notes: Values in parentheses are the prior means for (dim 1, dim 2). Prior variance is $0.01 \cdot \mathbf{I}_2$ *for anchors. Denmark (DNK) anchors dim 1 in 5 of 6 domains, reflecting its consistently high treaty ratification across domains. The dim-2 anchor is domain-specific and chosen to capture the primary confounding dimension (e.g., China for investment captures South-South BIT activity; the USA for human rights captures selective ratification of optional protocols).*

Estimation

The model is estimated via EM with Albert–Chib data augmentation (Albert and Chib 1993) and a Kalman filter-smoother for the dynamic ideal points. In the E-step, latent utilities y^*_{ij} are drawn from truncated normals conditional on current parameter estimates, and ideal points are updated via a forward-backward Kalman pass. In the M-step, item parameters (α_j, β_j) are updated via weighted least squares. Convergence is assessed by monitoring the log-likelihood, with a patience-based stopping rule (5 consecutive iterations with relative change below 10^{-2}). All models converge in 16–24 iterations. Five domains are estimated over $T = 7$ five-year periods (1990–2024); the security domain uses $T = 6$ periods (1990–2018), constrained by the temporal coverage of the ATOP alliance dataset (Leeds et al. 2002).

Hierarchical Joint Model

The per-domain IRT models estimate each domain independently, precluding direct estimation of cross-domain correlations or a unified latent scale. To address this limitation, we complement the per-domain analysis with a hierarchical joint IRT model that decomposes each country's position into a common factor capturing aggregate ILO support and domain-specific deviations.

Specifically, we model country i's latent position in domain d at period t as:

$$\theta_{i,d,t} = \mu_{i,t} + \gamma_{i,d}$$

where $\mu_{i,t}$ is a common factor representing overall ILO support and $\gamma_{i,d}$ is a time-invariant domain-specific deviation. The observation model is:

$$P(y_{ij} = 1 \mid \theta_{i,d,t}) = \text{logit}^{-1}(\alpha_j + \beta_j \cdot \theta_{i,d,t})$$

where α_j is the item difficulty, β_j is a scalar discrimination parameter, and the logistic link replaces the probit used in the per-domain models for computational convenience.

The common factor evolves as a random walk:

$$\mu_{i,1} \sim N(0, 1), \qquad \mu_{i,t} \mid \mu_{i,t-1} \sim N(\mu_{i,t-1},\, \omega^2)$$

with $\omega = 0.25$ fixed. We incorporate V-Dem's liberal democracy index as a soft predictor of μ via an additional prior:

$$\mu_{i,t} \sim N(b_v \cdot \widetilde{\text{VDem}}_{i,t},\, \sigma^2_{\text{reg}})$$

where $\widetilde{\text{VDem}}_{i,t}$ is the standardized V-Dem score (mean 0, SD 1) and $\sigma_{\text{reg}} = 1.0$ is fixed. The coefficient b_v is estimated and captures the extent to which domestic liberal democracy predicts aggregate ILO support. This is additive with the random walk: periods with observed V-Dem data contribute both the dynamic prior from the previous period and the V-Dem predictor.

The domain-specific deviations use a non-centered parameterization with optional predictors:

$$\gamma_{i,d} = \delta_d \cdot Z_{i,d} + \sigma_d \cdot \tilde{\gamma}_{i,d}, \qquad \tilde{\gamma}_{i,d} \sim N(0,1)$$

where $Z_{i,d}$ is a binary country-level predictor for domain d. We include two theoretically motivated predictors: tax haven status for investment (32 of 165 countries) and nuclear weapon state status for arms control (7 of 165). Human rights and intellectual property have no domain predictor ($Z_{i,d} = 0$). The parameter δ_d captures the predictor's effect on domain-specific deviation, while σ_d captures residual heterogeneity.

Identification is achieved by fixing two anchor countries across all periods: the USA at $\mu = +2$ (as the architect of the ILO) and Cuba at $\mu = -2$. The model is estimated via Hamiltonian Monte Carlo (HMC) in Stan (Carpenter et al. 2017), using 4 chains with 1,500 warmup and 2,000 sampling iterations each. The hierarchical model covers four domains — investment, human rights, arms control, and intellectual property — for 165 countries over seven periods ($N_{\text{obs}} = 70{,}715$). Environment and security are excluded: environment because its very large item set ($J = 3{,}111$) creates computational infeasibility, and security because its six-period temporal structure (ending in 2018) would create a ragged data structure.

Results

Model Fit

Table 4 reports the percent correctly classified (PCC) for each domain. PCC is computed as $\hat{y}_{ij} = \text{sign}(\alpha_j + \beta_j^{'}\mathbf{x}_{i,s(j)})$ over all non-missing entries.

Table 4. Classification accuracy (PCC) by domain and dimension

Domain	dim-1	dim-2	Both	Baseline	n
Investment	98.8%	98.5%	98.8%	98.5%	505,778
Security	97.6%	97.6%	97.8%	97.4%	94,001
Environment	94.6%	92.2%	95.8%	89.9%	481,335
Human Rights	87.9%	87.9%	89.8%	84.2%	11,171
Arms Control	84.6%	86.2%	86.8%	81.3%	5,769
Intellectual Property	92.5%	92.0%	93.5%	91.8%	16,151

Notes: Baseline = proportion of the modal category. dim-1 and dim-2 use only the respective dimension; Both uses the full two-dimensional model.

The model improves upon the modal baseline in all domains, with the largest gains in environment (+5.9 pp), human rights (+5.6 pp), and arms control (+5.5 pp). The second dimension contributes most in arms control (+2.2 pp over dim 1 alone) and human rights (+1.9 pp), consistent with the expectation that these domains contain confounding structure beyond unidimensional ILO support. By contrast, investment shows a PCC of 98.8% against a baseline of 98.5%, a gain of only 0.3 pp — reflecting the extreme sparsity of the bilateral investment treaty matrix, where the vast majority of country-treaty pairs are zeros and the modal category dominates. The high baseline means that the IRT model adds little predictive power in investment, even as it recovers substantively interpretable ideal points.

Ideal Point Estimates

Our quantity of interest is not any individual country's trajectory but rather the aggregate distribution of latent positions across periods, which captures systemic-level shifts in support for the ILO.

Investment

The investment domain (Figure 1) shows a steady rightward shift in the distribution of dim-1 ideal points from 1990–1994 (mean = 1.03) to 2020–2024 (mean = 1.55), reflecting the expansion of the bilateral investment treaty network. The trend continues into the most recent period, though the rate of increase slows markedly after 2005–2009 as the BIT network approaches saturation. The spread also increases (sd from 1.78 to 2.16), indicating growing heterogeneity. Tax havens and offshore financial centers tend to cluster in the high-dim-1 region, consistent with their role as active BIT signatories.

Security

The security domain (Figure 2), the only domain limited to six periods (1990–2018) due to the ATOP dataset's coverage, exhibits a leftward drift in dim-1 means (from -0.78 to -1.10), suggesting that the distribution of alliance commitments has shifted away from the anchor-defined ILO-support pole. Nuclear powers occupy a range of positions along both dimensions, reflecting the diversity of their alliance portfolios.

Environment

The environment domain (Figure 3), with the largest item set ($J = 3{,}111$), shows relatively stable dim-1 means (ranging from -0.66 to -0.61 across all seven periods through 2020–2024) but a clear downward shift in dim-2 means (from -0.68 to -1.18). The increasing spread on both dimensions indicates growing polarization in environmental treaty engagement. The 2020–2024 period continues this pattern without notable change.

Human Rights

In the human rights domain (Figure 4), the distribution shows a slightly negative dim-1 mean (around -0.15 to -0.21 across all seven periods through 2020–2024), with gradually increasing dispersion (sd from 1.02 to 1.34). This pattern reflects the near-universal ratification of core human rights treaties combined with selective ratification of optional protocols and newer instruments. The 2020–2024 period shows no notable change from the preceding period. The USA, anchoring dim-2, is positioned distinctly from countries that have ratified the full suite.

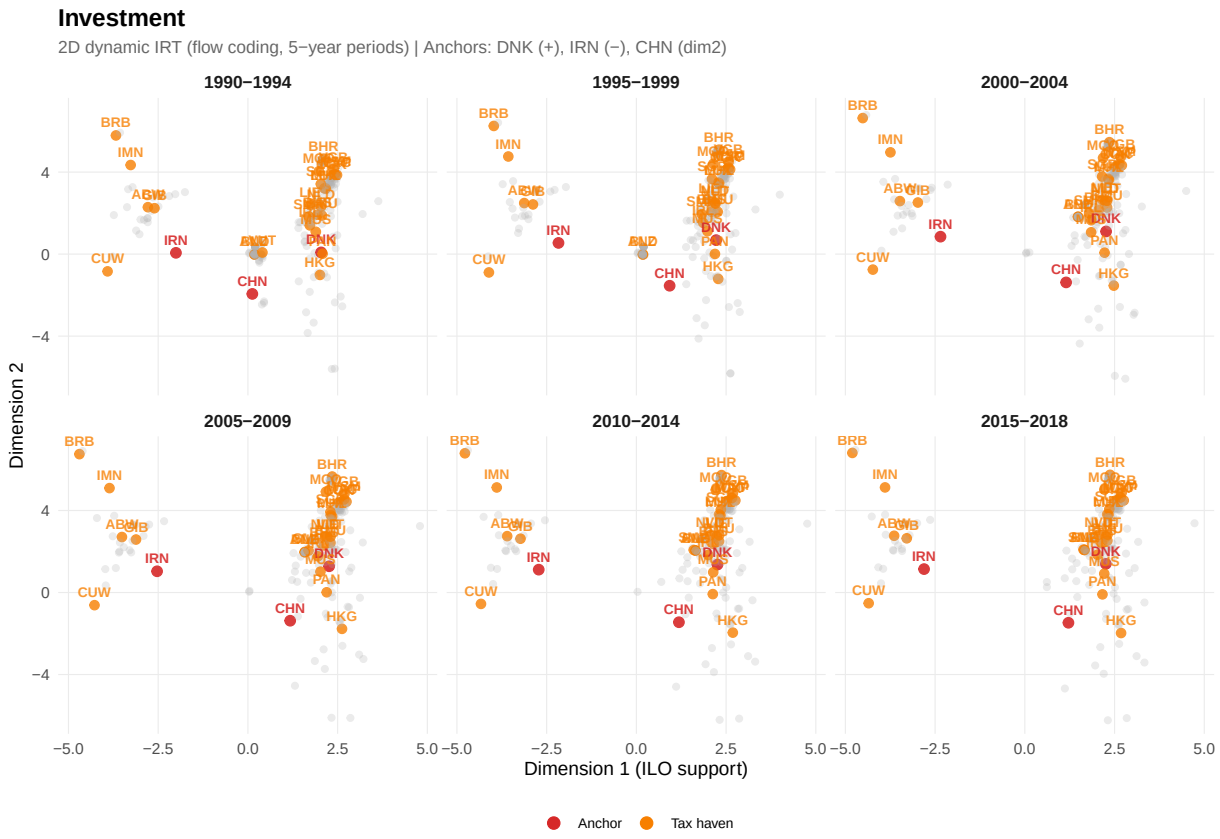

Figure 1: Investment ideal points (2D), 1990–2024. Anchors in red; tax havens in orange.

Figure 2: Security ideal points (2D), 1990–2018. Anchors in red; nuclear powers in blue.

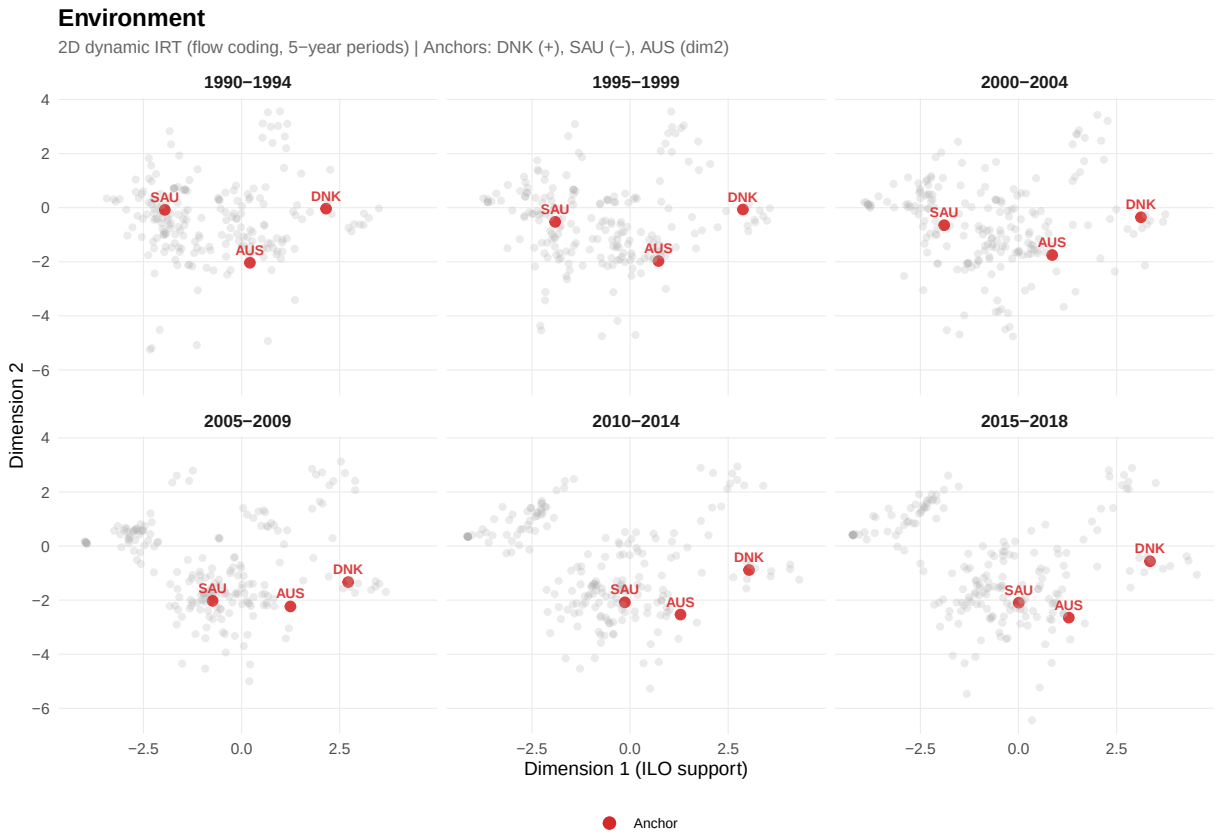

Figure 3: Environment ideal points (2D), 1990–2024. Anchors in red.

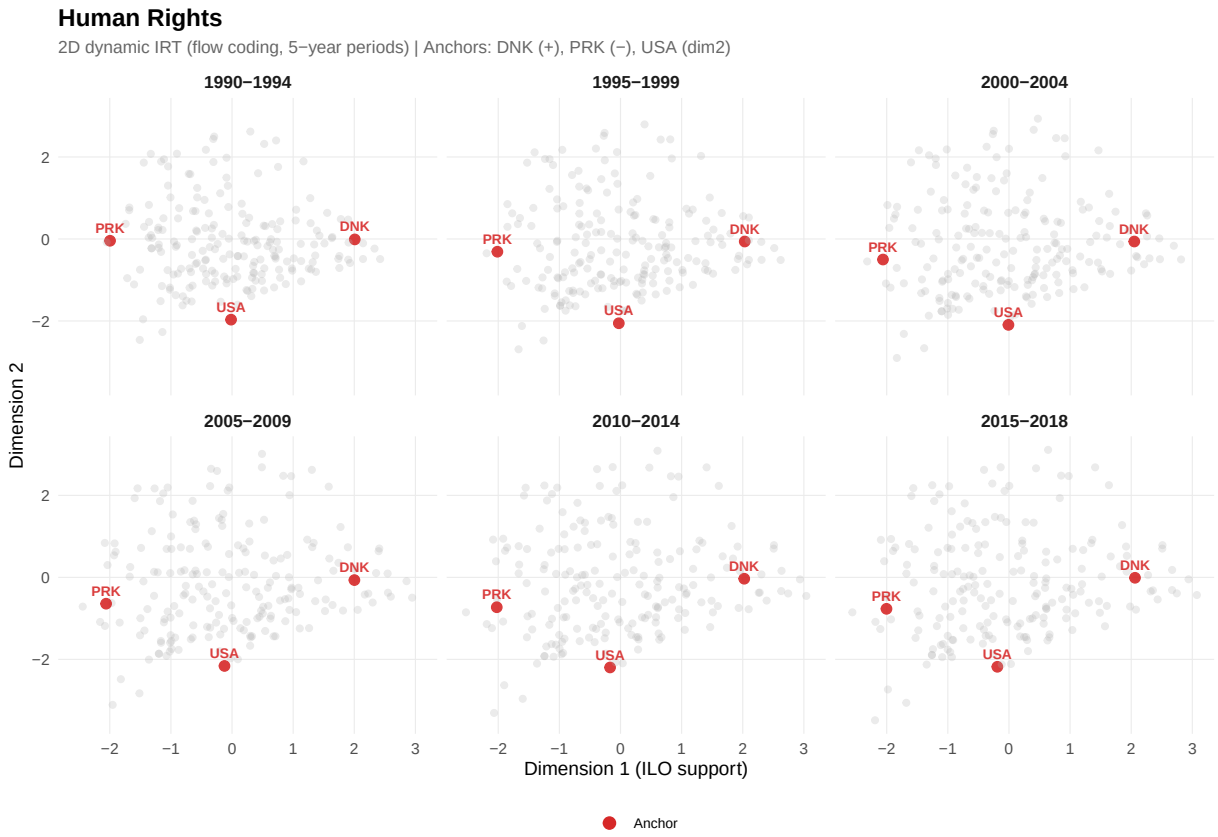

Figure 4: Human Rights ideal points (2D), 1990–2024. Anchors in red.

Arms Control

The arms control domain (Figure 5) includes 11 treaties, from the NPT (191 parties) to the Treaty on the Prohibition of Nuclear Weapons (69 parties). Dim-1 means hover near zero throughout the study period (from +0.10 in 1990–1994 to −0.01 in 2020–2024), but the distribution compresses from 1990–1994 (sd = 0.94) to 1995–1999 (sd = 0.91) before slowly widening to 1.09 by 2020–2024. The initial compression coincides with the wave of post-Cold War accessions to the CWC, CTBT, and Ottawa Convention.

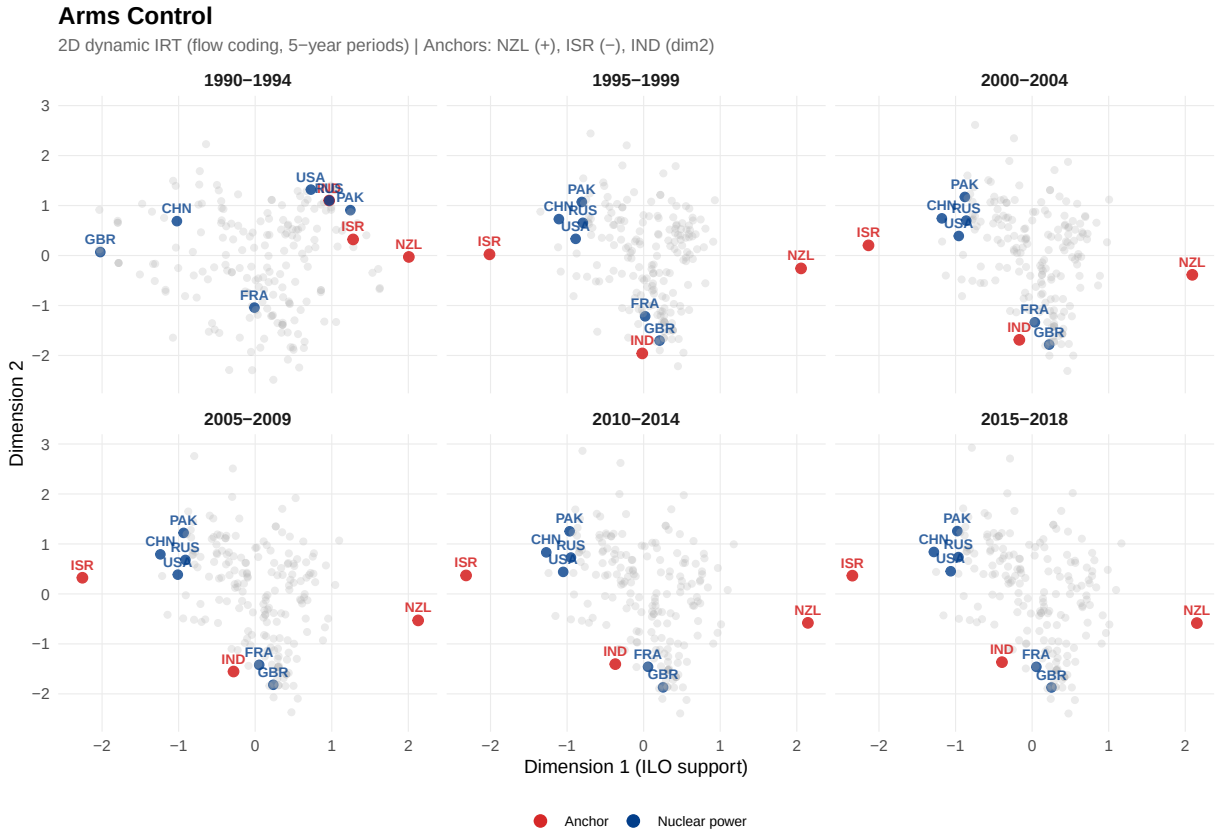

Figure 5: Arms Control ideal points (2D), 1990–2024. Anchors in red; nuclear powers in blue.

Intellectual Property

The intellectual property domain (Figure 6) shows a leftward drift in dim-1 means (from −0.06 in 1990–1994 to −0.26 in 2020–2024), reflecting the rapid expansion of WIPO treaty membership after the TRIPS Agreement. The drift plateaus after 2010–2014, suggesting the IP regime has reached a stable equilibrium. The spread increases substantially (sd from 0.99 to 1.34), indicating growing differentiation as some countries adopt the full suite of WIPO-administered treaties while others remain selective.

Trade

Unlike the six treaty-based domains, trade is measured directly through tariff data from the WITS/TRAINS database, covering 176 countries from 1990 to 2022. We construct two sum-

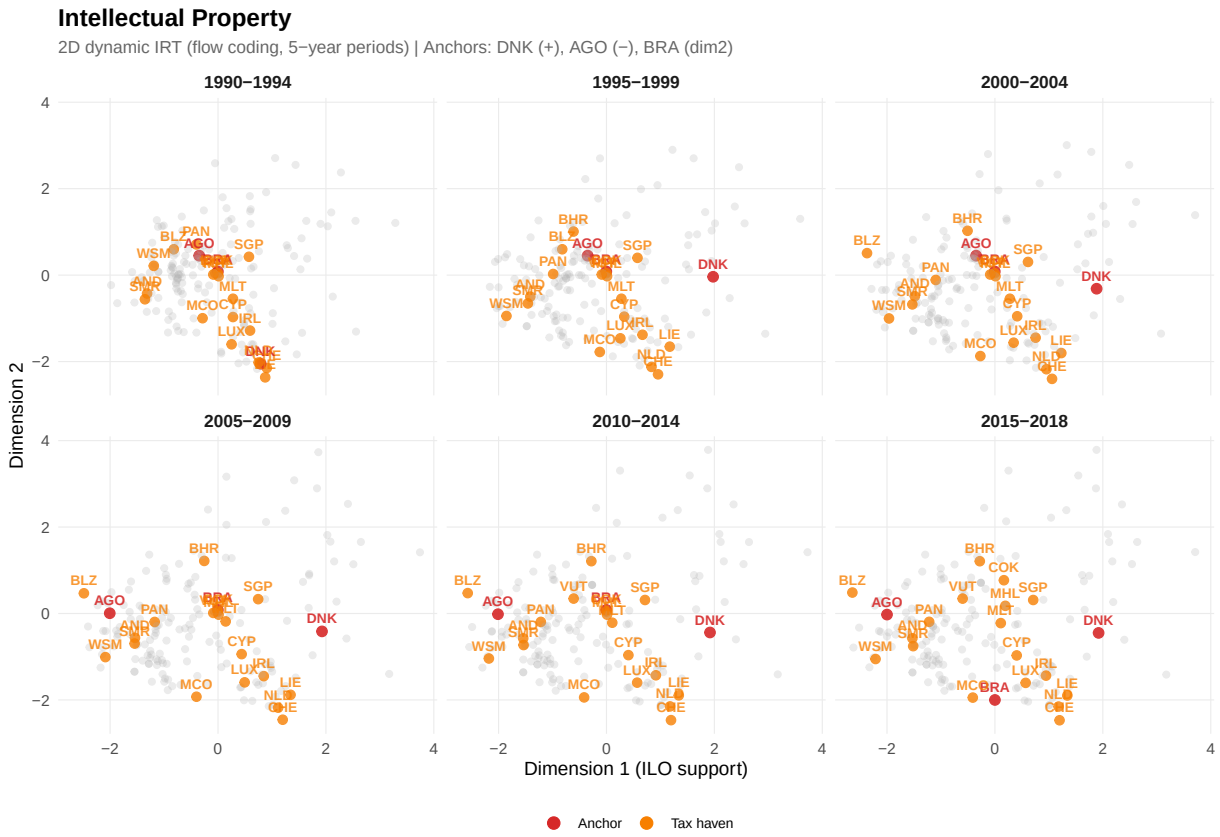

Figure 6: Intellectual Property ideal points (2D), 1990–2024. Anchors in red; tax havens in orange.

mary measures for each country-year: (1) the mean MFN applied tariff, averaged across 21 HS sections, and (2) the mean preferential margin, defined as the difference between the MFN tariff and the best preferential tariff available to any partner, averaged across sections. Together, these two measures distinguish three ideal types of trade engagement: *universally open* countries (low MFN, low margin), *club-open* countries (high MFN, high margin — open to preferential partners but not to the world), and *closed* countries (high MFN, low margin).

Figure 7 shows the evolution of these two measures. The cross-country mean MFN tariff falls steadily from 18.4% in 1990–1994 to 10.1% in 2019–2022, reflecting three decades of multilateral and unilateral liberalization. The cross-country dispersion also compresses (sd from 16.2% to 5.3%), indicating convergence toward lower tariffs. Meanwhile, the mean preferential margin rises from 8.8% to 12.3%, reflecting the proliferation of preferential trade agreements. The dispersion of margins *increases* (sd from 6.4% to 7.8%), indicating growing differentiation in preferential depth.

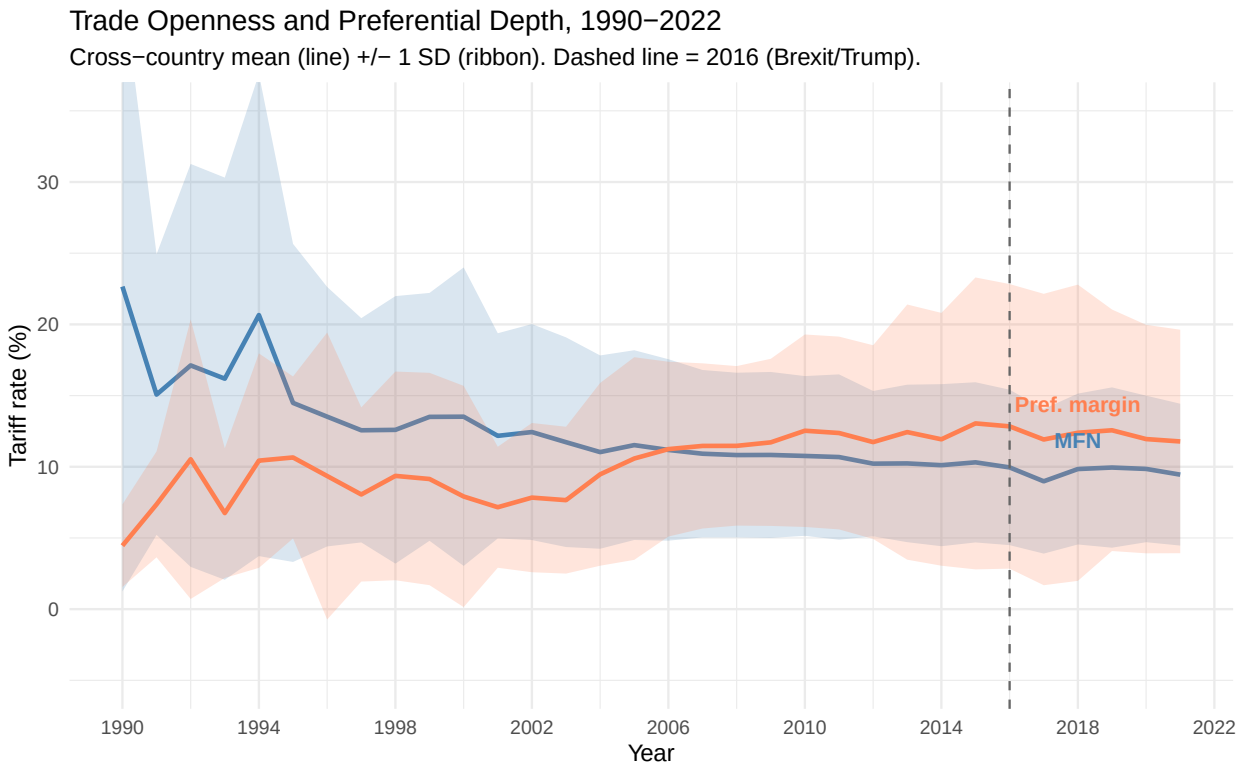

Figure 7: Trade openness and preferential depth, 1990–2022. Blue = MFN tariff; coral = preferential margin. Lines = cross-country mean; ribbons = ±1 SD. Dashed line = 2016 (Brexit referendum, Trump election).

Figure 8 plots MFN tariffs against preferential margins for the 2019–2022 period, revealing the three-type structure. The USA, Japan, Australia, and the UK cluster in the lower-left quadrant (universally open: low MFN, low margin). India, Brazil, South Korea, and Turkey occupy the upper-right (club-open: high MFN, high preferential depth through regional agreements like MERCOSUR and RCEP). Few countries occupy the lower-right (closed: high MFN, low margin), reflecting the near-universality of preferential agreements in the modern trade system.

Figure 9 shows MFN trajectories for selected countries. India and China exhibit the most dramatic liberalization (India from 68% to 14%; China from 41% to 8%), though India's tariffs have *risen*

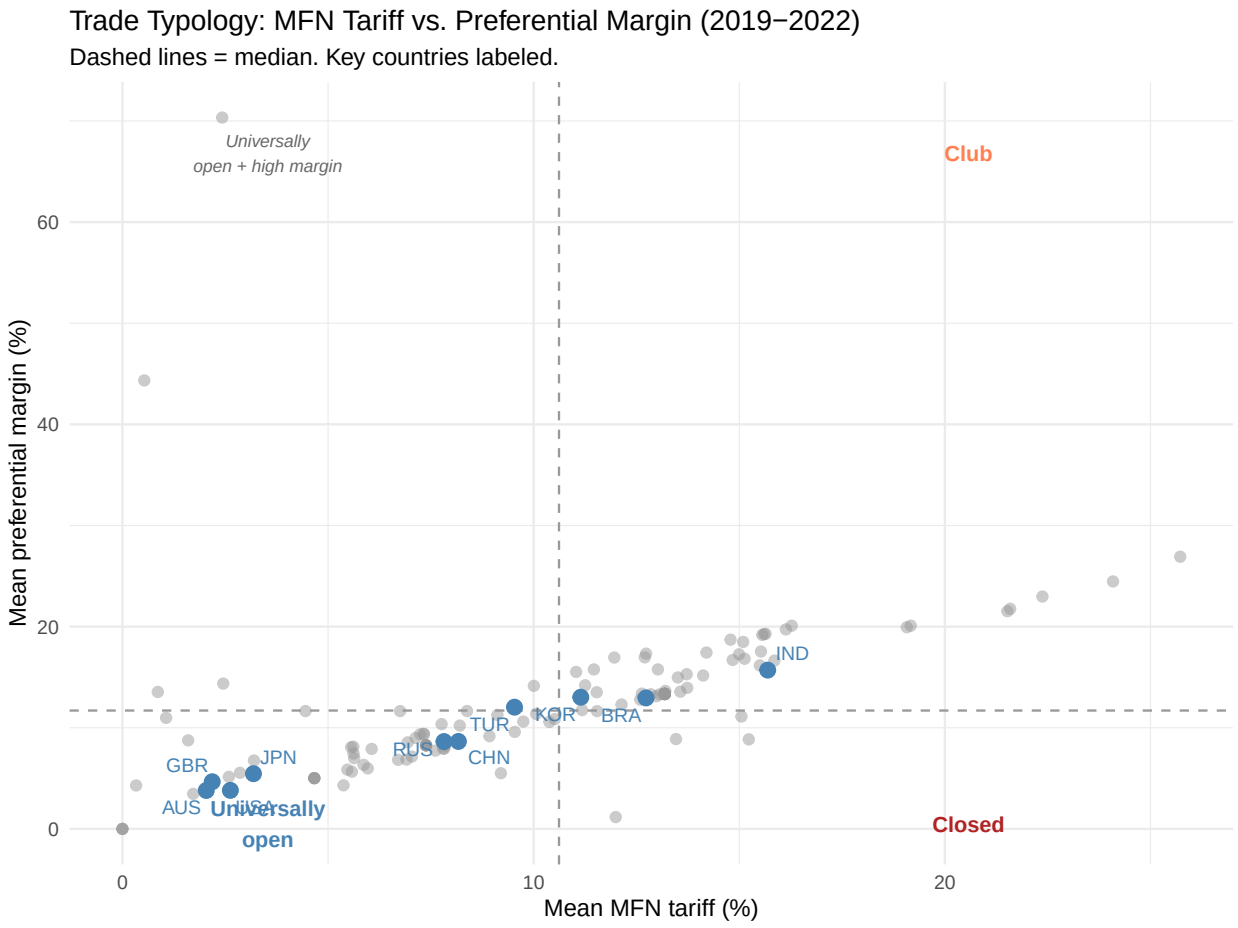

Figure 8: Trade typology: MFN tariff vs. preferential margin, 2019–2022. Dashed lines = cross-country medians. Key countries labeled.

since 2018 — one of the few cases of sustained MFN reversal in the dataset. The 2016 mark (Brexit/Trump) does not produce a visible aggregate break in MFN tariffs, likely because tariff increases under Trump were targeted (specific products and partners) rather than across-the-board.

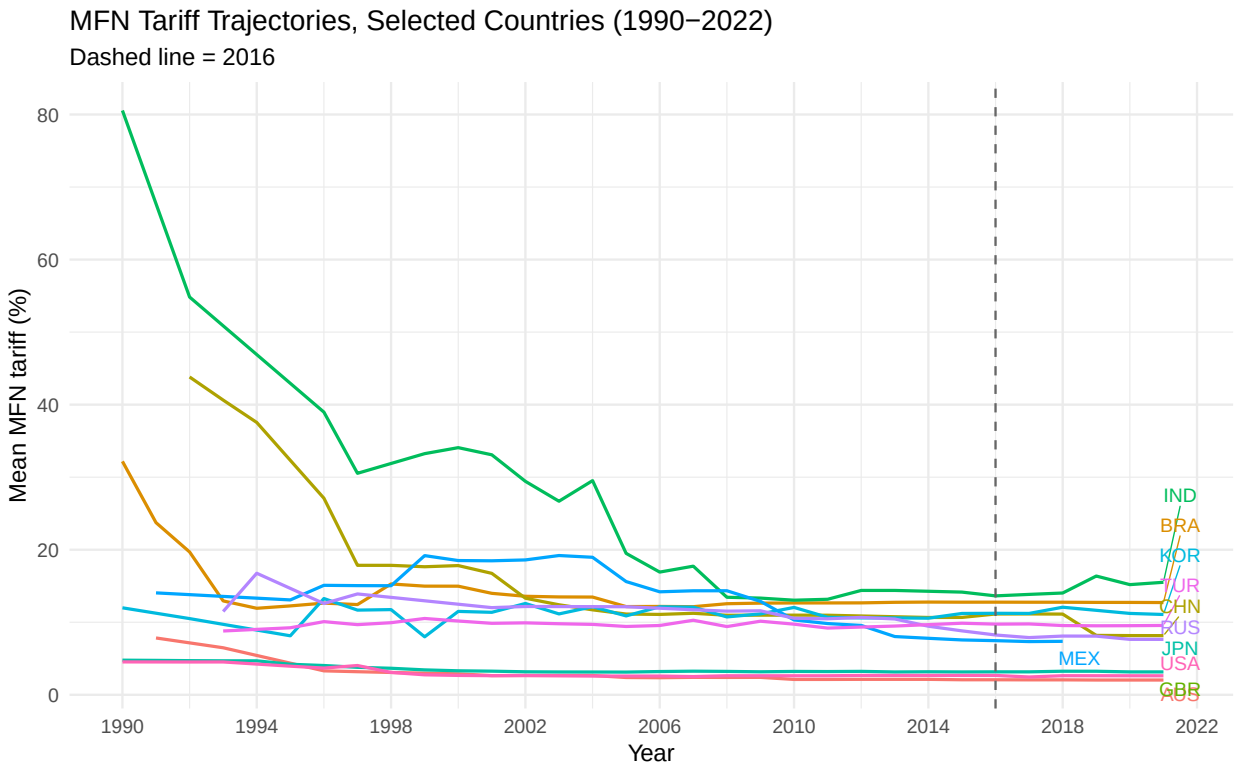

Figure 9: MFN tariff trajectories, selected countries, 1990–2022. Dashed line = 2016.

Aggregate Trends

Table 5 summarizes the temporal evolution of dim-1 ideal points across all six domains.

Table 5. Aggregate trends in dim-1 (anchor-defined ILO-support dimension), 1990–1994 vs. latest period

Domain	Mean (1990–94)	SD (1990–94)	Mean (latest)	SD (latest)	Latest period	Trend
Investment	+1.03	1.78	+1.55	2.16	2020–24	↗
Security	−0.78	1.68	−1.10	1.84	2015–18	↘
Environment	−0.66	1.43	−0.61	2.04	2020–24	≈
Human Rights	−0.15	1.02	−0.21	1.34	2020–24	≈
Arms Control	+0.10	0.94	−0.01	1.09	2020–24	≈
Intellectual Property	−0.06	0.99	−0.26	1.34	2020–24	↘
Trade (MFN tariff)	18.4%	16.2%	10.1%	5.3%	2019–22	↗ (liberalization)

Domain	Mean (1990–94)	SD (1990–94)	Mean (latest)	SD (latest)	Latest period	Trend
Trade (Pref. margin)	8.8%	6.4%	12.3%	7.8%	2019–22	↗ (more preferences)

Notes: Trend arrows: ↗ = increasing ILO support, ↘ = decreasing, ≈ = approximately stable. Five treaty domains extend through 2020–2024; security is constrained to 2015–2018 by ATOP dataset coverage. Trade rows report raw tariff rates (not latent ideal points); the trade period comparison uses 1990–1994 vs. 2019–2022. Values for the five extended domains are from the V8 estimation ($T = 7$ periods).

Three broad patterns emerge. First, the investment domain shows the clearest evidence of expanding treaty engagement along the anchor-defined ILO-support dimension, driven by the proliferation of bilateral investment treaties throughout the 1990s and 2000s; the extension through 2020–2024 confirms that this trend continues, though at a decelerating pace. Second, security shows a drift away from the pro-ratification pole, potentially reflecting the fragmentation of alliance structures after the Cold War (this domain cannot be assessed beyond 2018 due to data limitations). Third, the remaining domains show relatively stable means but increasing dispersion through the most recent period, suggesting that while the average level of treaty engagement has remained roughly constant, countries have become more differentiated in their commitment patterns.

The most striking finding is that the standard deviation of dim-1 ideal points increases over time in all six domains through the latest available period. This growing heterogeneity suggests that the ILO is not experiencing uniform erosion or uniform deepening, but rather increasing differentiation — a pattern consistent with the "contested multilateralism" thesis of Börzel and Zürn (2021). The 2020–2024 period, which covers the COVID-19 pandemic and intensified US-China competition, does not show a dramatic break from pre-existing trends in any of the five extended domains.

An important caveat is that these temporal trends are sensitive to the choice of coding scheme. Under stock coding (Section R1), where countries retain a +1 for all periods after initial ratification, temporal trends reverse in four of six domains (security, arms control, IP, and human rights). Stock coding produces a different picture because early adopters accumulate high scores over time, mechanically compressing the distribution as latecomers catch up. The growing-heterogeneity finding therefore reflects increasing differentiation in *new commitments* (flow coding) rather than in cumulative treaty portfolios (stock coding). Both perspectives are substantively meaningful, but readers should note that the headline finding is coding-dependent.

External Validation

To assess whether our treaty-based ideal points capture meaningful variation, we correlate dim-1 estimates with two external measures: UNGA ideal points (Bailey, Strezhnev, and Voeten 2017) and V-Dem's liberal democracy index (Coppedge et al. 2023). Table 6 reports pooled Pearson correlations across all country-period observations.

Table 6. External validation: correlation with UNGA and V-Dem

Domain	r (UNGA)	r (V-Dem)	N
Environment	0.48	**0.64**	1,315
Intellectual Property	**0.50**	0.33	1,264
Security	0.43	**0.52**	969
Human Rights	0.35	**0.36**	1,322
Arms Control	0.30	**0.50**	1,308
Investment	0.09	0.03	1,185

Notes: Bold indicates the stronger external correlate. Correlations computed over all non-missing country-period pairs. Five extended domains cover seven periods (1990–2024); security covers six periods (1990–2018).

Five of six domains show moderate-to-strong UNGA alignment ($|r| > 0.3$). V-Dem correlates more strongly than UNGA in four domains (environment, security, arms control, and human rights), suggesting that dim-1 partially captures domestic liberalism alongside international alignment. Intellectual property is the exception: UNGA outperforms V-Dem ($r = 0.50$ vs. 0.33), consistent with IP treaty participation tracking geopolitical alignment more than regime type. Investment shows near-zero correlation with both, reflecting the economic rather than political logic of BIT networks. Human rights shows a moderate positive correlation with both UNGA ($r = 0.35$) and V-Dem ($r = 0.36$), consistent with dim-1 capturing meaningful variation in ILO support in this domain. However, the anchor sensitivity documented in Section R2 (r as low as -0.87 with alternative anchors) indicates that identification remains fragile, and the human rights results should be interpreted with some caution.

Hierarchical Joint Model

The hierarchical joint IRT model estimates a common ILO-support factor μ alongside domain-specific deviations γ for four domains (investment, human rights, arms control, and intellectual property). The model converges with excellent diagnostics: zero divergences, maximum $\widehat{R} = 1.010$ (all parameters below the 1.01 threshold), and minimum effective sample size of 999 (median 3,663).

Table 7 reports the key parameter estimates.

Table 7. Hierarchical joint IRT: parameter estimates

Parameter	Estimate (SE)	Interpretation
b_v (V-Dem predictor)	0.352 (0.053)	Liberal democracy predicts ILO support (> 6 SE from zero)
$\delta_{\text{investment}}$ (tax haven)	1.08 (0.53)	Tax havens deviate positively in investment (~ 2 SE)
$\delta_{\text{arms control}}$ (nuclear)	-1.02 (0.37)	Nuclear states deviate negatively in arms control (~ 3 SE)
σ_d (investment)	3.34 (0.21)	Very high domain-specific heterogeneity
σ_d (human rights)	0.75 (0.09)	Moderate heterogeneity
σ_d (arms control)	0.60 (0.11)	Moderate heterogeneity
σ_d (intellectual property)	1.15 (0.13)	Moderate–high heterogeneity

Three findings emerge. First, the V-Dem coefficient $b_v = 0.352$ is well identified and positive, confirming that domestic liberal democracy is a meaningful predictor of aggregate ILO support as measured through treaty ratification. Countries with higher V-Dem scores tend to have higher μ, even after accounting for domain-specific deviations.

Second, the domain predictors have the expected signs. Nuclear weapon states deviate negatively in arms control ($\delta = -1.02$, approximately 3 standard errors from zero), consistent with the well-documented reluctance of nuclear powers to join disarmament treaties such as the TPNW. Tax havens and offshore financial centers deviate positively in investment ($\delta = 1.08$, approximately 2 standard errors), reflecting their active participation in the bilateral investment treaty network.

Third, the domain-specific variance σ_d reveals substantial heterogeneity in the degree to which domains are explained by the common factor. Investment is the most idiosyncratic ($\sigma_d = 3.34$), meaning that country positions in investment are driven primarily by domain-specific factors rather than general ILO support. This is consistent with the near-zero UNGA and V-Dem correlations reported in Table 6 and with the economic (rather than political) logic of BIT networks. By contrast, human rights ($\sigma_d = 0.75$) and arms control ($\sigma_d = 0.60$) are more tightly linked to the common factor, suggesting that treaty engagement in these domains is partially captured by a country's general orientation toward the liberal order. We can express this as a pooling factor $\omega^2/(\omega^2 + \sigma_d^2)$: arms control pools most strongly toward the common factor (0.15), followed by human rights (0.10) and IP (0.05), while investment is nearly independent (0.006).

The common factor μ produces substantively sensible rankings. The top positions (after the USA anchor at +2) are held by Slovenia (0.83), Montenegro (0.80), Spain (0.76), and Germany (0.74) — all countries with extensive treaty portfolios across multiple domains. The bottom positions are occupied by small Pacific island states (Marshall Islands, Tonga, Kiribati, Tuvalu) and authoritarian regimes (Myanmar, Eritrea) that have ratified few treaties. The common factor correlates at $r = 0.587$ with UNGA ideal points ($n = 162$), comparable to the per-domain correlations.

The domain-specific $\theta_{i,d}$ estimates from the hierarchical model correlate strongly with the per-domain dim-1 ideal points from the independent IRT models: $r = 0.85$ for investment, $r = 0.94$ for human rights, $r = 0.47$ for arms control, and $r = 0.84$ for intellectual property. The high correlations for investment, HR, and IP confirm that the hierarchical decomposition preserves the domain-level signal while additionally recovering the common factor. The lower correlation for arms control ($r = 0.47$) reflects the fact that the hierarchical model pools arms control most strongly toward the common factor, which alters country rankings relative to the independent model.

Robustness

We conduct five robustness checks to assess the sensitivity of our results to modeling choices. We summarize each briefly; full results are available in the supplementary materials.

R1: Stock Coding

Under stock coding, a country is coded as a party (+1) in all periods after initial ratification, rather than only in the period of ratification. This addresses the concern that flow coding creates an "activity bias" whereby long-standing members of treaties appear less committed in later periods because they have exhausted the treaty space.

Stock coding produces robust results for investment (trivially, since most BITs are single-period

events) and environment ($r > 0.96$ across all periods). However, it reverses temporal trends in four of six domains (security, arms control, IP, and human rights). This finding validates the activity bias concern: under stock coding, early adopters accumulate high scores, masking meaningful temporal variation in new commitments. We retain flow coding as the baseline because it better captures the dynamic signal of ongoing engagement, while reporting stock results for completeness.

R2: Alternative Country Anchors

We re-estimate each domain with two alternative sets of anchor countries, drawn from PCA extremes and theoretically motivated alternatives. Investment is highly robust ($r > 0.99$ across both alternative sets). Environment and intellectual property are robust ($r > 0.93$). Security is mostly robust ($r > 0.95$ for one set, $r = 0.96$ for the other). Human rights and arms control show greater sensitivity (r as low as -0.87 for HR with one alternative set), reflecting the challenge of identifying a consistent ILO-support dimension in domains with high treaty universality. Cross-sectional rankings are more stable than temporal trends.

R3: Item Anchor Sensitivity

We verify that theoretically motivated anchor items (e.g., NPT and CTBT for arms control, ICCPR for human rights) have the expected positive discrimination signs. In environment, arms control, and IP, all theory items have positive β_1 in the baseline model. Human rights shows a sign mismatch for one theory item, consistent with the identification challenges noted above. Constraining item start values to have positive β_1 produces ideal points that correlate at $r > 0.98$ with the baseline in five of six domains, with security as the outlier ($r = 0.79$).

R4: Evolution Variance (ω^2) Sensitivity

We vary ω^2 across a 50-fold range (0.01 to 0.5). All domains maintain $r > 0.97$ with the baseline for the moderate range ($\omega^2 \in [0.01, 0.2]$), and $r > 0.80$ even at the extreme $\omega^2 = 0.5$. Trend directions are preserved across all values. We conclude that $\omega^2 = 0.1$ is a safe default that does not drive substantive conclusions.

R5: Three-Year Temporal Windows

We re-estimate all models using 10 three-year periods (1990–1992 through 2017–2018, the last being a truncated two-year window) instead of 6 five-year periods. Investment, security, and IP show strong trend preservation ($r > 0.97$). Environment is moderately robust (trend $r = 0.82$). Human rights shows essentially zero trend correlation ($r = 0.08$), reflecting the near-zero trend in both specifications. Arms control shows a trend reversal ($r = -0.61$), likely driven by the extreme sparsity of the domain ($J = 38$ items under three-year windows) combined with the sensitivity of the first-period estimates. The sparse domains (HR, arms control) are most sensitive to temporal resolution, while the data-rich domains remain stable.

Discussion

Our results speak to three themes in the literature on the liberal international order.

Not simply decline, but differentiation. The dominant narrative frames the ILO as being in decline or crisis (Maull 2018; Cooley and Nexon 2020). Our results, at least under flow coding,

paint a more nuanced picture across three and a half decades through 2024. Only the security domain shows a clear drift away from the anchor-defined ILO-support pole (though this assessment is limited to the 1990–2018 window). Investment shows deepening that continues into 2020–2024. The remaining domains show stable means but increasing dispersion. Notably, the 2020–2024 period — marked by the COVID-19 pandemic and intensified geopolitical rivalry — does not produce a visible break from pre-existing trends in any of the five extended domains. This pattern is more consistent with Börzel and Zürn's (2021) framework of "contestation" — the order is not collapsing but becoming more heterogeneous, with growing gaps between countries that embrace the full institutional architecture and those that selectively engage.

Domain-specific dynamics. The ILO is not a monolith. Investment treaty networks follow economic logic (near-zero correlation with UNGA and V-Dem), security alliances and human rights track geopolitical alignment and domestic liberalism, and environmental agreements correlate strongly with both. These domain-specific dynamics are obscured in unidimensional measures. As Farrell and Newman (2021) argue, different institutional domains of the liberal order can evolve in contradictory directions, and our domain-specific measurement strategy is designed to capture precisely this heterogeneity.

Measurement as contested ground. Human rights ideal points correlate positively with both UNGA votes and liberal democracy scores ($r = 0.35$ and $r = 0.36$, respectively), suggesting that the anchor-defined first dimension does capture meaningful variation in ILO support. However, the high anchor sensitivity documented in Section R2 (r as low as -0.87 with alternative anchors) indicates that this positive alignment depends on the choice of anchor countries. As Loh, Paes, and Zarakol (2025) emphasize, what counts as "supporting" the liberal order depends on whose vision of the order one adopts. The HR results illustrate the challenge of imposing a single interpretive frame across domains with very different ratification structures: while the baseline anchors produce a dimension that aligns with external measures, alternative anchors can reverse this alignment entirely. We urge caution in treating the HR dim-1 scores as equivalent in meaning to dim-1 scores in other, more robustly identified domains.

Limitations. Several caveats apply. First, treaty ratification is an imperfect proxy for actual compliance. Countries may ratify treaties without implementing them, and some may comply without ratifying (Hathaway 2002; Simmons 2009). This ratification-compliance gap may be especially acute in the human rights domain, where ratification can serve as a low-cost signal of commitment without corresponding behavioral change (Kelley and Simmons 2015). Our measure therefore captures formal institutional engagement rather than substantive compliance with liberal norms. Second, our flow coding privileges recent activity, potentially underweighting the commitments of long-standing participants. Stock coding addresses this concern but introduces its own biases. Third, the per-domain estimation strategy estimates each domain's IRT model independently, which limits direct comparability across domains. The hierarchical joint model (Section 3.3) partially addresses this limitation by decomposing country positions into a common ILO-support factor and domain-specific deviations for four of six domains. However, this hierarchical model covers only investment, human rights, arms control, and intellectual property; environment is excluded due to computational constraints and security due to temporal mismatch. The hierarchical model also uses a one-dimensional common factor (rather than the two-dimensional specification of the per-domain models) and a logistic link (rather than probit), so results are not directly comparable. Fourth, the trade domain uses directly observed tariff rates (in percentage points) rather than the IRT framework applied to the other six domains. The IRT models produce latent ideal points on an arbitrary, anchor-normalized scale. The two types of measures are fundamentally incommensurate:

tariff rates have a natural unit and cardinal interpretation, while IRT ideal points are identified only up to the anchor normalization. Comparing trends across treaty-based and trade domains (as in Table 5) therefore involves comparing quantities on different scales, and aggregate statements about "heterogeneity across all domains" should be read as qualitative summaries rather than formally comparable estimates. Fifth, the human rights and arms control domains show identification sensitivity in several robustness checks (alternative anchors, item anchors, and temporal resolution), suggesting that the two-dimensional model may not fully capture the complexity of ratification patterns in these domains. Although the HR domain's external validation correlations are now positive ($r = 0.35$ with UNGA, $r = 0.36$ with V-Dem), the high anchor sensitivity (r as low as -0.87 with alternative anchors) raises questions about the robustness of this alignment. Arms control faces analogous challenges due to extreme item sparsity ($J = 49$ treaty-period items) and near-universal ratification of core treaties, which compress the latent space and make identification fragile. Results for these two domains should be interpreted with greater caution than for the data-rich domains (investment, environment).

Conclusion

We have proposed and implemented a measurement strategy for support for the International Liberal Order based on treaty ratification across six domains and tariff data for a seventh. Using a two-dimensional dynamic IRT model for the treaty domains and directly observed tariff measures for trade, we recover latent ideal points and policy indicators that capture systemic-level shifts in institutional commitment from 1990 to 2024. Five treaty domains now extend through the 2020–2024 period, allowing us to assess whether the turbulence of the early 2020s — the COVID-19 pandemic, intensified great-power competition, and rising economic nationalism — produced a visible break in treaty engagement patterns. Our central finding — increasing heterogeneity in new commitments across most domains under flow coding — persists into the most recent period, offering a quantitative counterpoint to both optimistic and pessimistic narratives about the order's trajectory, though this pattern is sensitive to coding choices (see Section R1). The trade domain reinforces this picture: while MFN tariffs have converged downward, preferential margins have diverged, reflecting a shift from multilateral to club-based liberalization. The evidence suggests that the ILO is not simply declining; it is differentiating, though the strength of this conclusion varies across domains and measurement choices.

References

Albert, James H., and Siddhartha Chib. 1993. "Bayesian Analysis of Binary and Polychotomous Response Data." *Journal of the American Statistical Association* 88(422): 669–679.

Bailey, Michael A., Anton Strezhnev, and Erik Voeten. 2017. "Estimating Dynamic State Preferences from United Nations Voting Data." *Journal of Conflict Resolution* 61(2): 430–456.

Börzel, Tanja A., and Michael Zürn. 2021. "Contestations of the Liberal International Order: From Liberal Multilateralism to Postnational Liberalism." *International Organization* 75(2): 282–305.

Carpenter, Bob, Andrew Gelman, Matthew D. Hoffman, Daniel Lee, Ben Goodrich, Michael Betancourt, Marcus Brubaker, Jiqiang Guo, Peter Li, and Allen Riddell. 2017. "Stan: A Probabilistic Programming Language." *Journal of Statistical Software* 76(1): 1–32.

Cooley, Alexander, and Daniel Nexon. 2020. *Exit from Hegemony: The Unraveling of the American*

Global Order. Oxford University Press.

Coppedge, Michael, et al. 2023. *V-Dem Codebook v13*. Varieties of Democracy (V-Dem) Project.

Destradi, Sandra, and Johannes Vüllers. 2024. "Populism and the Liberal International Order: An Analysis of UN Voting Patterns." *The Review of International Organizations* 19: 1–28.

Fariss, Christopher J. 2014. "Respect for Human Rights Has Improved over Time: Modeling the Changing Standard of Accountability." *American Political Science Review* 108(2): 297–318.

Farrell, Henry, and Abraham L. Newman. 2021. "The Janus Face of the Liberal International Information Order: When Global Institutions Are Self-Undermining." *International Organization* 75(2): 333–358.

Hathaway, Oona A. 2002. "Do Human Rights Treaties Make a Difference?" *Yale Law Journal* 111(8): 1935–2042.

Imai, Kosuke, James Lo, and Jonathan Olmsted. 2016. "Fast Estimation of Ideal Points with Massive Data." *American Political Science Review* 110(4): 631–656.

Kelley, Judith G., and Beth A. Simmons. 2015. "Politics by Number: Indicators as Social Pressure in International Relations." *American Journal of Political Science* 59(1): 55–70.

Leeds, Brett Ashley, Jeffrey M. Ritter, Sara McLaughlin Mitchell, and Andrew G. Long. 2002. "Alliance Treaty Obligations and Provisions, 1815–1944." *International Interactions* 28(3): 237–260.

Loh, Dylan M. H., Lucas De Oliveira Paes, and Ayşe Zarakol. 2025. "Non-Western Visions of International Order." *Annual Review of Political Science*.

Martin, Andrew D., and Kevin M. Quinn. 2002. "Dynamic Ideal Point Estimation via Markov Chain Monte Carlo for the U.S. Supreme Court, 1953–1999." *Political Analysis* 10(2): 134–153.

Maull, Hanns W., ed. 2018. *The Rise and Decline of the Post-Cold War International Order*. Oxford University Press.

Mitchell, Ronald B., et al. 2020. "What We Know (and Could Know) About International Environmental Agreements." *Global Environmental Politics* 20(1): 103–121.

Simmons, Beth A. 2009. *Mobilizing for Human Rights: International Law in Domestic Politics*. Cambridge University Press.

Voeten, Erik. 2021. *Ideology and International Institutions*. Princeton University Press.