Mapping the Global Catholic Educational Presence: Structural Configurations of National Education Systems

Isabella Cordisco

Pontifical Salesian University

Correspondence: Isabella Cordisco (cordisco.info@gmail.com)

ABSTRACT

The Catholic Church is among the largest non-state providers of education in the world, yet its presence remains largely invisible in macro-comparative analyses of national education systems. Integrating international statistical sources (UNESCO, World Bank) with data from the Annuarium Statisticum Ecclesiae for 165 countries, I apply principal component analysis and fuzzy cluster analysis to thirteen indicators of institutional capacity, participation, and demographic demand. I identify five recurring configurations of national education systems, distinguished by demographic pressure and degree of consolidation. Catholic educational presence reaches its maximum neither in the most fragile nor in the most consolidated systems, but in expansive systems under strong demographic pressure, where it is also markedly primary-school-heavy. Across the remaining configurations, its internal primary-to-secondary balance is heterogeneous rather than tracking a single gradient of consolidation. These findings are consistent with a logic of differentiated complementarity, rather than substitution or residuality, and illustrate the value of a configurational reading of religious presence in contemporary institutional fields.

Keywords: Catholic education | comparative education | cluster analysis | principal component analysis | religious institutions | configurational analysis

Replication Statement: *The dataset (integrating UNESCO, World Bank, and Annuarium Statisticum Ecclesiae sources) and analysis code used in this study are available from the author upon reasonable request.*

1 Introduction

The Catholic Church operates one of the most extensive non-state networks of educational institutions in the world, present in nearly every national school system and in contexts ranging from consolidated welfare states to countries with fragile public infrastructures. Yet this presence constitutes a blind spot in comparative research. Macro-level analyses of education systems do not systematically include religious actors among their analytical dimensions, while the rich tradition of scholarship on Catholic education has developed within frameworks devoted to identity, mission, and ethos rather than to large-scale systematic comparison (Grace and O'Keefe 2007; D'Agostino and Carozza 2019). The result is that we know surprisingly little, at the comparative level, about a structurally significant component of many education systems. In particular, we do not know where it is concentrated relative to the different configurations of educational demand and system capacity.

This question is relevant to the social scientific study of religion as much as to comparative education. A well-established line of research has shown that religious actors operate as central providers of welfare and social services, and that the meaning of this presence is never legible in the abstract. It depends on the relationship with the state, the welfare regime, and the institutional context in which religious presence is embedded

(Bäckström and Davie 2010; Bäckström et al. 2011; Göçmen 2013). This reading is part of the broader sociological thesis that religious institutions have not withdrawn into the private sphere but continue to occupy a public role in contemporary institutional fields, a thesis formulated by Casanova (1994) and subsequently extended to a global, comparative scale (Casanova 2019). If this is true for welfare, it should hold a fortiori for education, historically the most extensive component of churches' social presence. *The sociologically interesting question is therefore not how much Catholic education exists in the world, but what position it occupies within different national configurations* — and what this positioning reveals about the role religious institutions play in contemporary institutional fields.

The Catholic case is particularly well suited to this question because it cannot be reduced to a single function (Grace and O'Keefe 2007; Grace 2016). In some contexts, Catholic schools have supported the extension of basic education where public coverage was limited (Wodon 2014; Grace and Wodon 2021). In others, they have developed as an integrated component of subsidiary or mixed arrangements, under public financing and state regulation (Bäckström et al. 2011; Wodon 2019). In still others, they occupy distinctive segments, associated with dynamics of school choice and institutional differentiation (Dronkers and Robert 2008; Dronkers and Avram 2010).

Catholic presence cannot therefore be interpreted a priori either as a substitute for state shortfalls or as a marginal residue of consolidated systems. It appears to take on different meanings depending on the configurations within which it is embedded. Yet this intuition, recurrent in the literature, has never been subjected to systematic empirical testing on a global scale. This intuition finds explicit theoretical grounding in James's (1993) model of the public-private mix in education, which distinguishes a component of excess demand, linked to insufficient public spending, from a component of differentiated demand, attributable to cultural and religious heterogeneity. These two components anticipate, respectively, the readings of substitution and differentiation developed in §2.4.

The present study takes this step. Integrating international statistical sources (UNESCO, World Bank) with data on the Catholic Church's educational activities drawn from the Annuarium Statisticum Ecclesiae, I construct a dataset of 165 countries. The analytical strategy proceeds in three steps. The first is a principal component analysis of thirteen indicators describing institutional capacity, participation and outcomes, and the demographic and social demand placed on education systems. The second is a fuzzy cluster analysis on the principal components. It identifies recurring configurations without imposing rigid memberships. The third examines how Catholic educational presence, primary and secondary, is distributed within these configurations. This last step includes a regression test of whether the distribution is attributable to the share of Catholic population, net of cluster membership.

The contribution of this paper is twofold. Empirically, the study reveals a systematically overlooked component of education systems, linking ecclesiastical data to international

statistical infrastructures on this scale for the first time. Theoretically, it proposes and tests a configurational reading of the relationship between religious presence, educational demand, and national system capacity, even as it distinguishes three rival interpretations of Catholic positioning — substitution, complementarity, differentiation — and their respective empirical expectations.

The article proceeds as follows. First, I situate the problem within the literature on religious actors as service providers and show why Catholic education constitutes its most macroscopic yet least comparatively studied case. Second, I present the configurational perspective guiding the analysis and the three readings of Catholic positioning. I then turn to the research questions, data and methods, results, and a discussion of the implications for the social scientific study of religion.

2 Background

2.1 Religious Actors and Service Provision in Institutional Fields

The sociology of religion has extensively documented that religious organizations are not confined to the cultic sphere but operate as structural providers of social services — welfare, health care, education — across all world regions. Comparative European research coordinated by Bäckström and Davie has shown that the role of churches in welfare varies systematically with the welfare regime and with the history of Church-State relations. It is marginal and complementary where the welfare state is pervasive. It is substantial and resurgent where the state retreats or never fully arrived (Bäckström and Davie 2010; Bäckström et al. 2011). Comparative studies of faith-based organizations confirm that the division of labor between the state and religious actors is not a given, but an institutional outcome that takes different forms across countries (Göçmen 2013; Cnaan 2002).

At the global scale, the quantitative literature has further suggested systematic relationships between public social spending and the vitality of religious presence, indicating that the two spheres move in relation to one another — of substitution, complementarity, or specialization — rather than independently (Gill and Lundsgaarde 2004). The religious economies paradigm has proposed reading the vitality and location of religious presence as the outcome of supply-side dynamics as much as demand-side ones (Finke and Stark 1992). At the micro level, analyzing National Congregations Study data, Chaves and Tsitsos (2001) consistently show that social services provided by religious congregations do not constitute a sector alternative to the secular one, but are an integral part of it.[1]

For the Catholic case, this interweaving also has a precise doctrinal matrix. The principle of subsidiarity assigns to intermediate bodies — families, communities, ecclesial institutions — a primary role in responding to needs, and reserves for the state a supporting

[1] See also Chaves et al. 2020 for the most recent wave.

and guaranteeing function (Pius XI 1931; Pontifical Council for Justice and Peace 2004). Subsidiarity is invoked here not as a normative category but as an analytical key. It posits that Catholic educational presence would be location-sensitive to national configurations of demand and capacity, rather than a uniform or purely demographic distribution. Surprisingly, however, empirical verification of this expectation has been lacking precisely in the field where Catholic presence is most extensive — education. The literature on religion and welfare has concentrated on assistance and social services, predominantly in Europe and North America. The global scale and the educational sphere remain largely unexplored.

2.2 Catholic Education and Its Macro-Comparative Invisibility

This gap does not stem from an absence of studies on Catholic education, which constitutes, on the contrary, a broad and vital field, but one developed through thematic perspectives and circumscribed comparisons. The reference literature, from the International Handbook of Catholic Education (Grace and O'Keefe 2007) to Sullivan's (2001) work on the tension between the distinctive and inclusive character of Catholic schooling, concentrate attention on identity, mission, leadership, and the relationship with secularization. Even contributions that situate Catholic schools in the face of contemporary challenges — inequality, cultural change, tensions between mission and market (Grace and O'Keefe 2007; Grace 2016) — keep the focus on educational mission rather than on the structural location of Catholic presence within national configurations of demand and supply.

Bryk, Lee, and Holland (1993) showed that Catholic schools can be read as specific institutional forms, endowed with distinctive organizational and communal properties. That insight invites a shift of the analysis from the identitarian to the structural plane. It has not, however, translated into a tradition of macro-comparative analysis. D'Agostino and Carozza (2019) explicitly note this gap, proposing a research agenda for the international and comparative study of Catholic education.

The line of research on the Catholic school effect was opened by Coleman, Hoffer, and Kilgore (1982) on the comparative achievement of public and Catholic schools. It likewise remains confined to the North American context. Its focus is on learning outcomes rather than on the structural location of Catholic presence within national education systems. A more recent review, edited by the same D'Agostino together with Asadullah (2025), confirms that this research agenda remains largely unmet on a global scale.

Correspondingly, international education statistics absorb Catholic presence into the general category of non-state actors (UNESCO 2021; Day Ashley et al. 2014), losing its historical and organizational specificity. Secular institutional networks, with their own systems of governance, teacher training, and community rootedness, are treated as an undifferentiated aggregate of private providers. The result is the double invisibility that this study aims to overcome: Catholic education is too “particular” for macro-comparison and too

"structural" for identity-based scholarship. Even the literature on the global governance of education, while acknowledging the growing weight of non-state actors, tends to treat them as an undifferentiated category rather than differentiating their institutional logics (Verger, Fontdevila, and Zancajo 2017).

This double invisibility is not a random artifact of data collection, but a predictable effect in light of world polity theory. On this account, international statistical infrastructures do not simply record what exists. They incorporate and diffuse a rationalized script of educational modernity in which the state is the legitimate actor by definition. Religious actors are relegated to the residual category of non-state actors, when they are not omitted altogether (Meyer, Boli, Thomas, and Ramirez 1997). From this perspective, the systematic absence of Catholic presence from the very sources that shape comparative education policy does not reflect a neutral measurement choice, but the very logic by which the global education field defines who counts as a legitimate education provider. Linking ecclesiastical data to international statistical infrastructures on this scale for the first time thus means not only filling an empirical gap, but rendering visible — and potentially renegotiating — a specific hierarchy of institutional legitimacy.

2.3 Beyond Additive Indicators: A Configurational Reading of Education Systems

Overcoming this impasse requires, first of all, an adequate way of representing national education systems. Comparative analysis has long relied on indicators of access, enrollment, resources, and performance, read in additive terms or as rankings (OECD 2024). Yet countries with similar levels of enrollment or spending can present profoundly different configurations in terms of demographic pressure, structural vulnerability, and the articulation of supply. A growing body of literature has therefore argued that the central question is not the expansion of schooling per se, but the overall coherence of systems — the capacity to hold together demand, resources, and outcomes (Pritchett 2015; Hanushek and Woessmann 2007). To this must be added the caution, arising from studies of the global governance of education, that international indicators are not neutral instruments but frame and define the very terms of comparison (Elfert and Ydesen 2023; Steiner-Khamsi et al. 2024).

From these premises follows the choice of a configurational reading. Such reading represents each national system not as a score on individual indicators, but as a combination of latent dimensions — demand pressure, vulnerability, capacity, participation — and identifies the recurring configurations in which such combinations cluster. It is within these configurations, rather than against one-dimensional rankings, that the positioning of Catholic presence becomes legible.

The choice of a configurational reading also finds a more properly historical-institutional motivation. National education systems are the outcome of path-dependent historical trajectories — the order and timing of the expansion of mass schooling, the nature

of Church-State relations inherited from the colonial or postcolonial period, the critical junctures at which public-private arrangements consolidated. This produces equifinality: countries with similar levels on a single indicator may share very little of the institutional history that produced them. An additive model, which treats variables as independent and substitutable effects, cannot distinguish between configurations that are historically irreducible to one another. A configurational analysis instead makes it possible to treat causation as conjunctural and multiple. This is the logic of multiple conjunctural causation that has inspired the development of set-theoretic methods (Ragin 2000, 2008).

2.4 Substitution, Complementarity, Differentiation: Three Readings of Catholic Positioning

Observed within structural configurations, Catholic educational presence can be interpreted according to three readings, each with distinct empirical expectations. These readings do not correspond to a typology already established within a single research tradition. They form an analytical construct that I propose here in order to render comparable a set of expectations developed so far within circumscribed and mutually unconnected bodies of work: studies of Catholic and faith-based schools in sub-Saharan Africa, comparative research on welfare and religion in Europe, and studies of school choice and institutional differentiation. James's (1993) model of the public-private distribution of education, which theoretically distinguishes a component of excess demand from a component of differentiated demand linked to cultural and religious heterogeneity, provides the scaffolding that allows me to bring these traditions together within a single space of rival hypotheses.

The first, which I term the reading of substitution (or compensation), predicts that Catholic presence concentrates where public coverage is weaker and structural vulnerability higher, thereby filling the gaps of state provision. Studies of Catholic and faith-based schools in sub-Saharan Africa document a significant role precisely in contexts where demand remains high and public supply faces greater difficulties (Wodon 2014; Grace and Wodon 2021). If this reading were generalizable, we should expect maximum Catholic presence in the most fragile configurations.

The second, which I term the reading of institutional complementarity, predicts that Catholic presence flourishes not where the state is absent, but where public systems are expanding and demand pressure exceeds consolidation capacity. In such settings it absorbs excess demand within subsidiary or mixed arrangements (Bäckström et al. 2011; Wodon 2019). In this case, the peak of presence should be located in intermediate, high-pressure configurations, rather than at the extremes of the distribution.

The third, which I term the reading of differentiation, predicts that in consolidated systems Catholic presence does not disappear but repositions itself in distinctive segments: secondary education, specialized provision, and school choice. These segments are

intertwined with the logics of family preference and institutional pluralism (Dronkers and Robert 2008; Dronkers and Avram 2010). Independent evidence supports this institutional logic beyond a single research tradition. Comparative work on school sectors shows that private and religious schools respond to market and institutional logics distinct from those of the public sector (Davies and Quirke 2007). Recent research on institutional differentiation among school sectors, including Catholic schools, confirms that this segmentation remains an active and contemporary pattern rather than a historically bounded one (Sciffer et al. 2026). The corresponding empirical expectation does not concern the overall volume of presence, but its internal composition — a profile relatively more oriented toward the secondary level in consolidated systems.

The reading of differentiation can further be anchored to an explicit mechanism drawn from organizational ecology. The niche width model developed by Hannan and Freeman (1977; Freeman and Hannan 1983) describes a structural trade-off between generalist and specialist strategies in exploiting environmental resources. As the carrying capacity of an organizational environment increases, non-state organizational forms tend to be pushed out of low-specialization segments. In the present case, that carrying capacity increases as public-system consolidation expands the available generalist supply. These forms are then driven to reposition themselves into narrower niches, where residual demand rewards distinctive rather than substitutive provision. Applied to the Catholic case, this mechanism predicts not a simple quantitative retreat in consolidated systems, but a qualitative shift in the internal profile toward segments — secondary education and specialized provision — that are less exposed to direct competition with the generalist public sector.

It is not possible here to isolate the specific contribution of other nonprofit and third-sector organizations relative to the non-state sector as a whole (see §4.1). The mechanism should therefore be understood as pertaining to non-state organizational forms in general. Catholic presence is the case observed here rather than a theoretically isolated exception.

To these three readings must be added an alternative explanation of a historical-demographic nature: Catholic educational presence might simply reflect the distribution of the Catholic population and the historical legacy of ecclesial settlement, independently of the configurations of education systems. This possibility has real grounding in the literature on the long-run effects of missionary settlement. Historical missionary presence has been shown to shape present-day educational and institutional outcomes largely independently of contemporary politics (Woodberry 2012). Comparative work on colonial-era Christian missions in Africa documents lasting effects on education specifically tied to inter-denominational competition for converts, a dynamic that could equally produce persistent cross-national variation in Catholic educational geography (Gallego and Woodberry 2010). The historical and regional heterogeneity in the market share of faith-inspired schools, tied to factors such as state subsidies or the historical expulsion of religious orders, further illustrates

how path-dependent this geography can be (D'Agostino and Asadullah 2025). In such cases, the distribution of presence across structural configurations should turn out to be weakly associated (or not at all) with system profiles.

The four readings are not mutually exclusive — they may operate simultaneously across different configurations — and it is precisely this possibility that calls for a configurational design, capable of observing at once where presence concentrates and what form it takes.

3 The Current Study

In this study I empirically reconstruct the structural configurations of national education systems and observe how Catholic educational presence is distributed within them. The analysis is guided by three specific questions.

First, where is Catholic educational presence concentrated? If the reading of substitution is correct, presence should reach its maximum in the most fragile, highest-vulnerability configurations. If complementarity prevails, the peak should be located in expanding configurations, where demand pressure exceeds consolidation capacity. If presence simply tracked systems' level of development, it should increase monotonically with their consolidation.

Second, is the distribution of Catholic presence independent of structural configurations, or does it adapt to them? A substantially uniform distribution across configurations would be consistent with the historical-demographic explanation, according to which presence responds to its own logics — religious demography, historical legacy — independent of system profiles. A structured distribution would instead suggest that Catholic presence is sensitive to configurations of demand and capacity.

Third, how does the internal profile — primary versus secondary — of Catholic presence vary across configurations? If the reading of differentiation is correct, consolidated systems should show a Catholic profile relatively more oriented toward the secondary level, while in high-pressure configurations presence should remain concentrated in basic education.

Informed by these three questions and by the rival readings discussed in §2.4, I formulate the following hypotheses:

H1. If the reading of substitution prevails, Catholic educational presence reaches its maximum in configurations with the highest structural vulnerability and demographic pressure (regardless of the system's level of consolidation).

H2. If the reading of institutional complementarity prevails, Catholic educational presence reaches its maximum in intermediate, expanding configurations, where demand pressure exceeds consolidation capacity — not at the extremes of the distribution.

H3. If the reading of differentiation prevails, the internal profile of Catholic presence (primary-to-secondary ratio) shifts progressively toward the secondary level as system consolidation increases, regardless of the overall volume of presence.

A fourth expectation concerns the historical-demographic explanation. It is formulated not as a hypothesis but as a control condition. If the distribution of Catholic presence across configurations were entirely attributable to the share of Catholic population, the differential between clusters should disappear once this control is introduced into the regression models.

4 Data and Methods

4.1 Data

The data derive from the integration of multiple international statistical sources — the UNESCO Institute for Statistics and the World Bank's World Development Indicators for indicators on education systems and demographic context — with data on the Catholic Church's educational activities drawn from the *Annuarium Statisticum Ecclesiae*. To my knowledge, this latter linkage constitutes the first attempt at this scale to connect ecclesiastical statistics to international statistical infrastructures. The unit of analysis is the country. The starting sample comprises countries with at least one indicator available from any of the three sources. After excluding countries lacking linkable ecclesiastical data, and territories not recognized as sovereign states in the UNESCO/World Bank sources, the final analytic sample comprises 165 countries.

For each variable, I selected the most recent year offering the broadest coverage across countries and used that single year's values for all countries for that variable, so as to maximize both temporal recency and international comparability. The reference year can therefore differ across variables (and across sources), though it is held constant across countries within each variable.

The share of missing values is limited for most indicators (0–19%), with the exception of the pupil-teacher ratios (56% in primary and 65% in secondary), for which imputation plays a substantial role. A sensitivity analysis re-running the full pipeline (multiple imputation via chained equations, PCA, fuzzy c-means) after excluding these two variables indicates reasonable stability. The two components together explain 60.7% of variance in the full model versus 59.2% when the two variables are excluded, and 89.7% of countries retain the same cluster assignment (optimal label matching via the Hungarian algorithm, since cluster indices are arbitrary across different runs).

A second check addresses imputation uncertainty more directly: the full pipeline (PCA and fuzzy clustering) was re-run separately on each of the five imputed datasets, and each was compared with the one used for the main analysis. Agreement in cluster assignment ranged from 84.8% to 86.7% across the four comparisons (mean 85.9%), indicating that the partition

is reasonably, though not perfectly, stable across imputations. Descriptive statistics for all variables included in the analysis are reported in Table 1.

Missing data on education-system indicators were handled through multiple imputation by chained equations (MICE), using predictive mean matching for all variables with missing values (m = 5 imputations, 40 iterations, seed fixed for reproducibility). The predictor matrix was specified selectively (minimum correlation 0.15; minimum proportion of usable cases 0.25). Two additional precautions were adopted. Variables on the share of teachers in the population were excluded from the predictor role. Feedback loops between pairs of strongly correlated variables were also interrupted (share of children and infant mortality; share of children and primary students per 1,000 inhabitants), so as to prevent collinearity from distorting the imputations.

Convergence was checked via trace plots, and consistency between original and imputed data was verified by comparing correlation matrices before and after imputation. Principal component analysis and fuzzy clustering were conducted on a single completed dataset (the first of the five imputations), since neither procedure has a standard extension for pooling results across multiply imputed datasets under Rubin's rules.

Two checks reported below in this section address the influence of this choice: a sensitivity analysis excluding the two variables with the highest share of missing values, and a direct comparison of the cluster partition obtained from each of the five imputed datasets. The regression reported in §5.5, by contrast, was re-estimated separately on each of the five imputed datasets, with results combined via Rubin's rules. This pooled estimate is the one reported as the primary result in §5.5. None of the variables entering that model is among the imputed indicators, so the between-imputation variance is null and the pooled estimate coincides exactly with the one obtained on a single completed dataset.

One informational limitation must be stated at the outset: the presence of secular, non-Catholic nonprofit education providers is not considered in the analysis owing to the lack of comparable data at the global scale. The implications of this exclusion are discussed in the concluding section.

Table 1. Descriptive Statistics of the Variables Included in the Analysis (N = 165 Countries)

Variable	N	Min	Max	Mean	SD	% Missing
Public spending on education (% of	153	0.00	8.44	4.10	1.62	7.3
Primary teachers (% of population)	148	0.0018	0.0124	0.0047	0.0017	10.3
Secondary teachers (% of population)	139	0.0007	0.0138	0.0060	0.0028	15.8
Pupil-teacher ratio, primary	73	8.88	59.51	25.89	12.47	55.8
Pupil-teacher ratio, secondary	57	7.34	72.31	18.07	10.87	65.5
Gross primary enrollment (%)	149	61.78	144.81	102.45	12.52	9.7
Gross secondary enrollment (%)	134	11.01	155.96	85.62	30.56	18.8

Variable	N	Min	Max	Mean	SD	% Missing
Primary completion rate (%)	140	40.76	116.31	88.88	15.52	15.2
Share of children in population (%)	165	10.51	49.00	26.36	10.45	0.0
Population growth rate (%)	165	−2.28	7.32	1.30	1.31	0.0
Primary students per 1,000 inhabitants	160	36.10	238.42	99.82	44.96	3.0
Secondary students per 1,000	151	26.75	139.06	80.84	22.87	8.5
Infant mortality (per 1,000 live births)	163	1.60	72.60	19.55	17.23	1.2
Catholic presence, primary (%)	161	0.00	95.90	6.92	12.11	2.4
Catholic presence, secondary (%)	153	0.00	82.88	5.51	10.31	7.3
Catholic presence, composite primary+secondary (%)	165	0.00	71.22	3.77	7.66	0.0

Note. Sources: UNESCO Institute for Statistics (education-system indicators); World Bank, World Development Indicators (demographic and contextual indicators); Annuarium Statisticum Ecclesiae (numerators for Catholic educational presence), with denominators from UNESCO/World Bank. N = 165 countries. For each variable the most recent year with the broadest country coverage was used; the reference year may therefore differ across variables. Percentages of missing values refer to the sample before multiple imputation.

4.2 Measures

Education system configurations. The analysis includes thirteen variables capturing three dimensions of national education systems. The first dimension concerns resources and institutional capacity: public spending on education as a percentage of GDP; primary and secondary school teachers relative to population; the pupil-teacher ratio in primary and secondary education. These variables describe the financial effort and staffing endowment with which systems respond to demand. The share of teachers in the population, in particular, measures the depth of a system's professional infrastructure, while pupil-teacher ratios signal its degree of strain.

The second dimension concerns participation and outcomes: gross primary and secondary enrollment rates and the primary completion rate. Together, these distinguish formal access from the capacity to retain students and bring them to completion. The third dimension concerns demographic and social demand: the share of children in the population, the population growth rate, primary and secondary students per 1,000 inhabitants, and infant mortality. The first four measure the volume and dynamics of the potential and actual demand weighing on systems. Infant mortality is included not as a health outcome, but as a synthetic indicator of the structural vulnerability of the context in which the education system operates — a choice that avoids introducing economic development variables (such as GDP per capita) whose relationship with education is notoriously bidirectional. The use of demographic indicators, rather than measures of expressed demand, further avoids circularity with the enrollment indicators placed on the participation side.

Catholic educational presence. The Church's educational presence is measured as the percentage share of students enrolled in Catholic educational institutions out of the total number of students at the corresponding level, calculated separately for the primary and secondary levels using numerators from the *Annuarium Statisticum Ecclesiae* and denominators from international statistical sources. A composite indicator synthesizes presence across the two combined levels.

4.3 Analytical Strategy

The analytical strategy unfolds in three steps, consistent with the configurational perspective outlined above.[2]

First, a principal component analysis was conducted on the thirteen standardized variables of the imputed dataset to identify the latent structural dimensions. The decision on the number of components to retain was based primarily on the scree plot, which shows a clear elbow after the second component. The Kaiser criterion (eigenvalue > 1) was treated as a secondary reference, since it is known to produce an excessive number of components in the presence of many correlated variables (Cattell 1966; Jolliffe 2002). The third component exceeds it only marginally (1.03), explaining a modest additional share of variance (7.9%) without adding an interpretively distinct dimension relative to the first two.

Throughout the paper I follow the common convention of referring to the standardized rotation coefficients from this analysis as "loadings." Formally, these are the eigenvector coefficients produced by the principal component decomposition, not loadings rescaled by the square root of the corresponding eigenvalue, and Table S2 (Supporting Information) is labeled accordingly.

Second, a fuzzy cluster analysis (fuzzy c-means algorithm; fuzzification parameter m = 2; Euclidean distance) was applied to the standardized scores of the retained components. Unlike hard clustering, fuzzy clustering assigns each country a degree of membership in each cluster, recognizing the existence of hybrid or transitional cases — a property substantively appropriate for national configurations that rarely fall neatly into discrete categories (Kaufman and Rousseeuw 2005). Solutions from k = 2 to k = 6 were evaluated using four internal validity indices (fuzzy partition coefficient, partition entropy, Xie-Beni index, and a fuzzy silhouette approximation).[3]

These indices do not point unanimously to the same solution: the partition coefficient, partition entropy, and fuzzy silhouette all favor coarser partitions and decline monotonically as k increases, a well-known tendency of these measures. The Xie-Beni index, which more

---

[2] For the combined use of dimension reduction and classification in constructing typologies, see OECD/European Union/JRC 2008. See also Kaufman and Rousseeuw 2005.

[3] See Table S3, Supporting Information.

directly balances within-cluster compactness against between-cluster separation, is instead minimized at k = 5 on the retained two-component space.

The five-cluster solution was therefore selected on the combined basis of this index, the substantive interpretability of the resulting profiles, and their stability across repeated initializations, rather than because every index converged on the same value of k. The two-cluster solution was retained as a macro-partition check for robustness analyses. For descriptive analyses, each country was assigned to the cluster with maximum membership, while membership degrees were retained as an indicator of hybridity. Across the 165 countries, the maximum membership degree ranges from 0.34 to 0.997 (median 0.74), with 27 countries (16.4%) below 0.5 and 46 (27.9%) below 0.6, confirming that a non-trivial share of cases are hybrid or transitional rather than sharply typed.[4]

Third, the distribution of Catholic educational presence — overall, primary, and secondary — was examined across clusters. This step draws on summary statistics (medians), on representations of the primary-secondary composition by cluster, and on a joint examination of Catholic presence, system consolidation (secondary enrollment), and structural vulnerability (infant mortality) at the cluster level. To test the historical-demographic explanation, presence was examined in two further ways. The first was relative, comparing the share of students enrolled in Catholic institutions to the share of Catholics in each national population. The second was a regression model of overall presence (log-transformed) on cluster membership, net of the Catholic population share. Given the nature of the design — ecological and cross-sectional — the analysis is descriptive-interpretive in character and does not advance causal claims.

5 Results

5.1 Two Structural Gradients

The principal component analysis indicates a clear latent structure. The scree plot shows a clear elbow after the second component, and the first two components are the only ones with eigenvalues above unity (6.23 and 1.67, respectively), jointly explaining 60.7% of total variance (PC1: 47.9%; PC2: 12.9%).[5] The third component has an eigenvalue only marginally above unity (1.03; 7.9% additional variance) but was not retained. The scree plot shows a clear elbow after PC2, and its loadings do not add a substantively distinct dimension relative to the demand-vulnerability/consolidation gradient and the schooling-intensity gradient already captured by PC1 and PC2.

The first component orders countries along a gradient from demographic pressure and structural vulnerability to educational consolidation. The highest positive loadings concern

[4] See Table S5, Supporting Information.

[5] Table S1 (Supporting Information); Figure S1 (Supporting Information): scree plot.

the primary pupil-teacher ratio (0.37), the share of children in the population (0.36), infant mortality (0.36), and primary students per 1,000 inhabitants (0.32). The strongest negative loadings concern secondary enrollment (−0.36), primary completion (−0.32), and the share of secondary teachers in the population (−0.27). The second component captures schooling intensity and system utilization, with dominant loadings for primary enrollment (−0.61), secondary students (−0.48) and primary students (−0.38) per 1,000 inhabitants, and the share of primary teachers (−0.33).[6] Given these raw loadings, higher PC1 values indicate greater demographic pressure and structural vulnerability rather than consolidation. Lower PC2 values indicate greater schooling intensity, since the relevant loadings are negative. No sign flip was applied to either component. The distribution of countries in the PC1–PC2 space appears structured and non-random, with consolidated European systems and high-pressure sub-Saharan African systems occupying opposite poles of the first gradient.[7]

5.2 Five Recurring Configurations

The fuzzy cluster analysis on the scores of the first two components identifies five recurring configurations (Figure 1: clusters in PC1–PC2 space; Figure 2: heatmap of structural profiles; Figure 3: world map). The standardized mean profiles allow for a clear interpretation. Cluster 3 (consolidated, low-pressure systems) combines low demographic pressure and vulnerability (share of children: −1.02; infant mortality: −0.85) with above-average primary completion and secondary enrollment (+0.52; +0.72). Most European countries, North America, and the consolidated systems of East Asia and Oceania fall in this category.

Cluster 2 (expansive primary systems under pressure) shows high pressure and vulnerability (share of children: +1.28; infant mortality: +1.19; pupil-teacher ratio: +1.48) together with the highest concentration of primary students (+1.81), against a backdrop of still-weak consolidation especially at the secondary level (secondary enrollment: −1.15). These are systems undergoing full expansion of basic education, whose demand exceeds consolidation capacity.

Cluster 4 (fragile, high-pressure systems) shows the most critical profile: high pressure and vulnerability (share of children: +1.39; infant mortality: +1.59) with the lowest values for primary completion (−2.00) and secondary enrollment (−1.56). Cluster 1 (moderately consolidated, transitional systems) shows contained pressure, medium-low vulnerability, and fairly solid educational levels (completion: +0.53; secondary enrollment: +0.41). Cluster 5 (intermediate or mixed systems) exhibits a profile close to the overall average, consistent with a hybrid or transitional configuration. It should be emphasized that, consistent with the fuzzy

[6] See Table S2 (Supporting Information).

[7] Figure S2 (Supporting Information): scatterplot of countries in PC1–PC2 space

approach, these clusters should not be read as rigid categories but as areas of clustering within a continuous structural space.

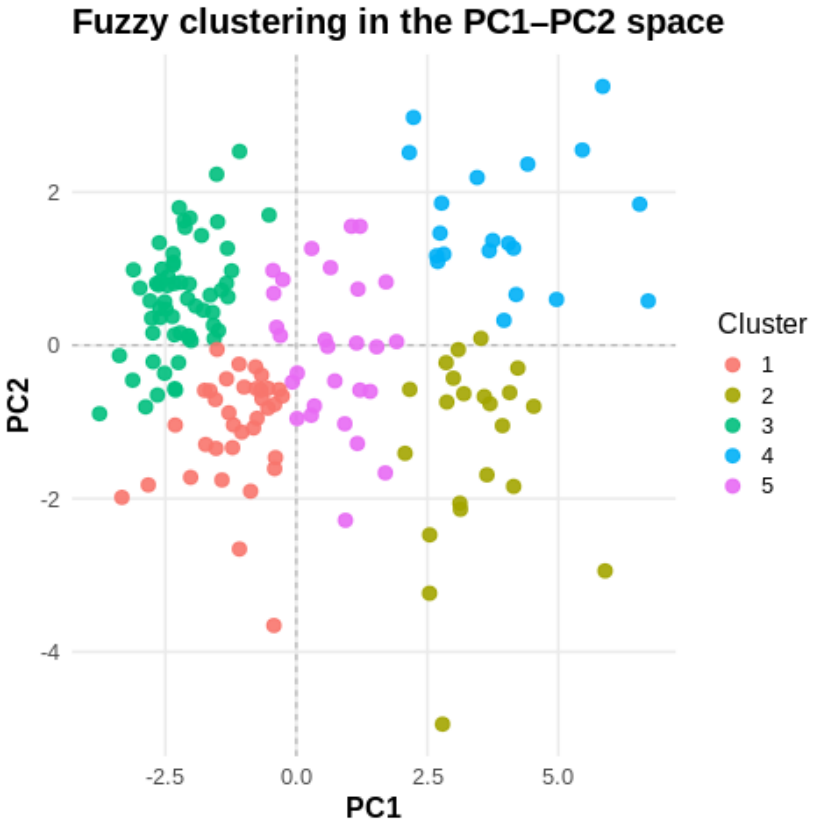


Figure 1. The Five Clusters in PC1–PC2 Space

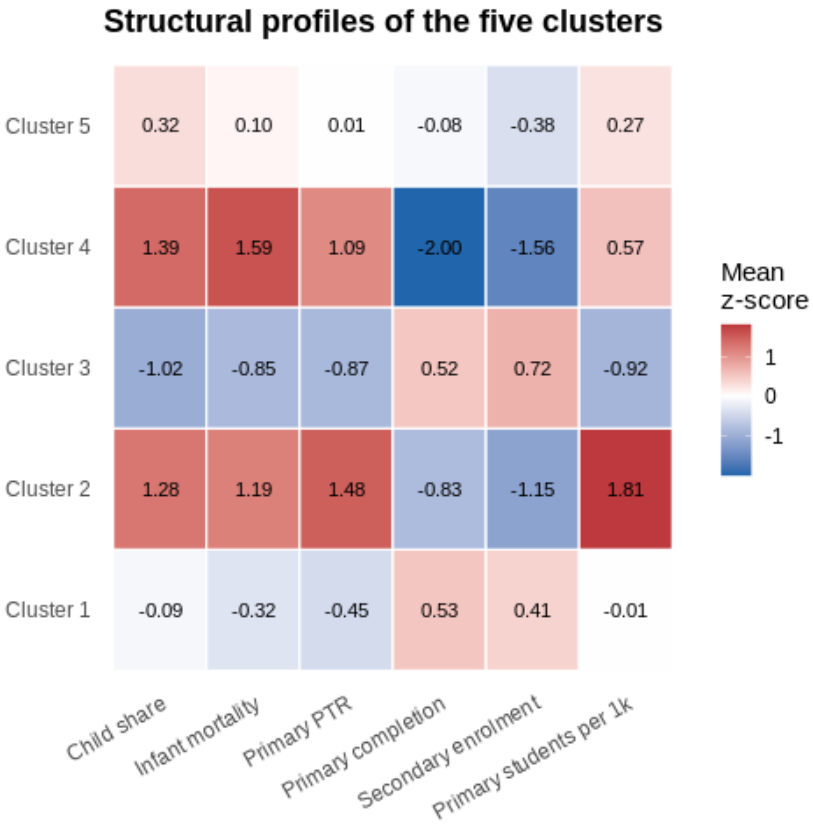


Figure 2. Heatmap of Standardized Structural Profiles by Cluster

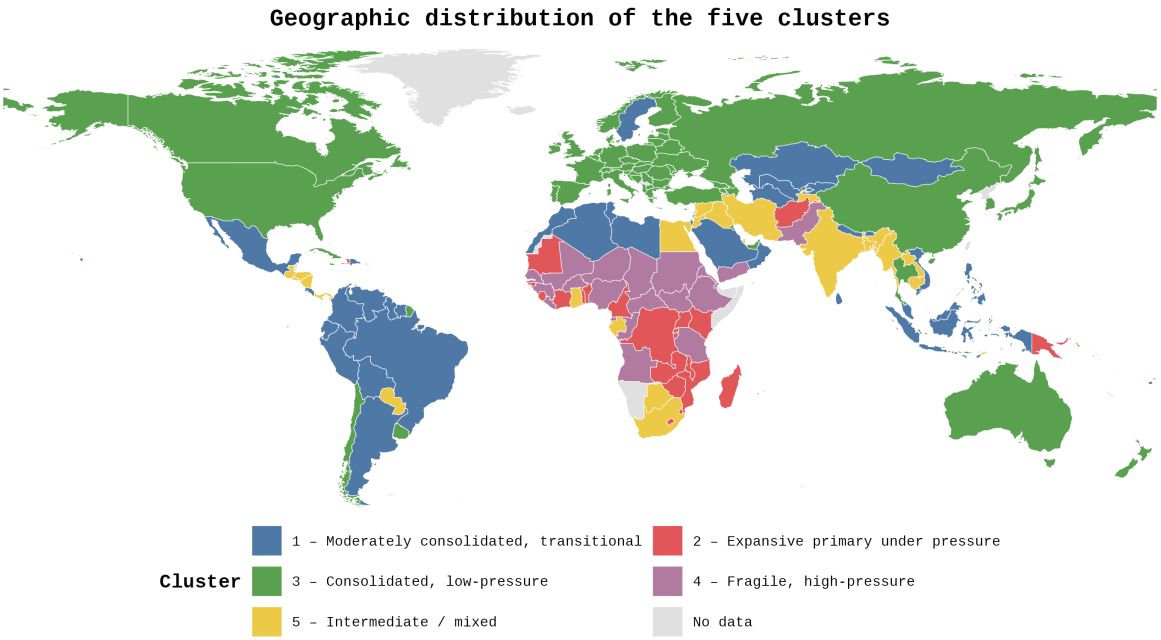


Figure 3. Geographic Distribution of the Five Clusters (numbering consistent with §5.2: Cluster 1 = moderately consolidated, transitional; Cluster 2 = expansive primary under pressure; Cluster 3 = consolidated, low-pressure; Cluster 4 = fragile, high-pressure; Cluster 5 = intermediate or mixed)

5.3 Where Catholic Presence Concentrates

Examining Catholic educational presence across configurations yields the study's first central result. Catholic presence reaches its maximum neither in the most consolidated nor in the most fragile configurations, but in the expansive-primary cluster under pressure (Cluster 2). There, median presence exceeds that of every other cluster on every measure: 12.7% of primary enrollees and 7.1% of secondary enrollees attend Catholic institutions, against 1.9% and 1.1% in consolidated systems (Cluster 3) and 2.0% and 2.7% in fragile systems (Cluster 4).[8] Figure 4 shows the underlying distributions: in every configuration the distribution is strongly right-skewed, with a minority of countries reaching very high values. Figure 5 positions the clusters jointly by system consolidation (secondary enrollment) and structural vulnerability (infant mortality), with Catholic presence represented by bubble size. The pattern is clear: the peak of presence lies in the intermediate, high-pressure quadrant and not at the extremes.

[8] Figure 4: Catholic presence by cluster

This result jointly answers the first two research questions. Against the reading of mechanical substitution, Catholic presence is not maximal where needs are most acute. In Cluster 4, the most fragile, it is present but at contained levels. Against the purely historical-demographic explanation, the distribution of presence is strongly structured with respect to configurations. If presence merely reflected religious demography and historical legacy, it should not cluster systematically in one specific configuration of pressure and capacity. Against the hypothesis that presence simply tracks systems' development, it does not grow monotonically with consolidation. In the consolidated systems of Cluster 3, overall presence is lower than in Cluster 2. The observed pattern is instead consistent with a logic of institutional complementarity: Catholic presence is more pronounced where public systems are expanding under a demand pressure that exceeds their consolidation capacity.

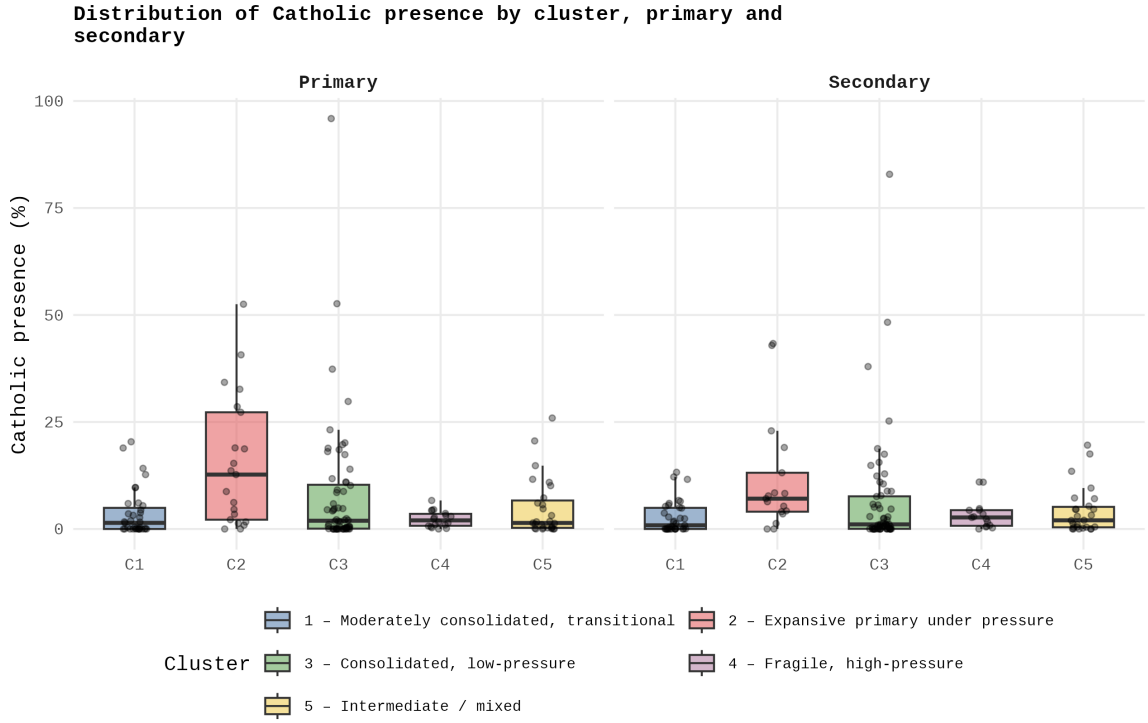


Figure 4. Distribution of Catholic Presence Across Countries, by Cluster and Level

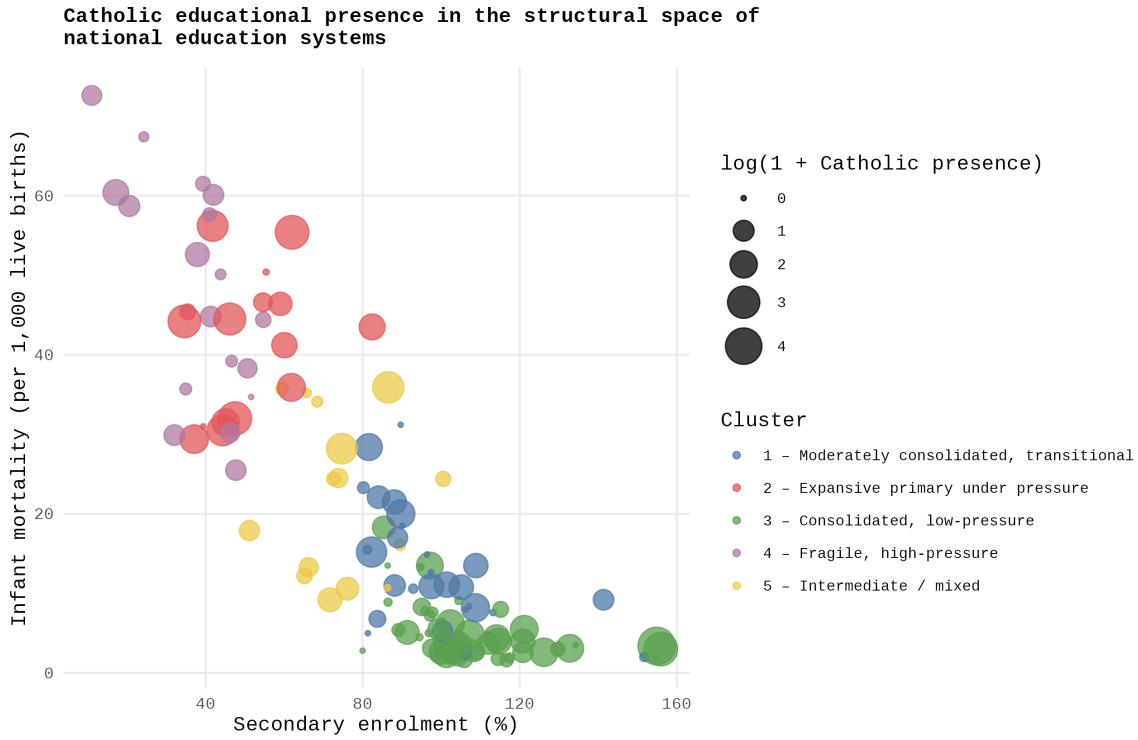


Figure 5. System Consolidation and Structural Vulnerability by Cluster (Bubble Size = Catholic Presence)

5.4 The Primary–Secondary Profile: Differentiated Complementarity

The internal composition of presence adds a second central result, though the pattern is more heterogeneous than a simple reversal. In Cluster 2, the Catholic footprint is skewed toward the primary level (medians: 12.7 vs. 7.1), consistent with a function of absorbing demand for basic education. In the fragile (Cluster 4: 2.0 vs. 2.7) and intermediate (Cluster 5: 1.4 vs. 2.0) configurations, the secondary component exceeds the primary one. Consolidated systems, however, do not follow this trend. Despite low overall volumes, their profile remains primary-heavy in relative terms (Cluster 3: 1.9 vs. 1.1), and the transitional systems are primary-heavy as well (Cluster 1: 1.4 vs. 0.9).[9] The internal primary-to-secondary ratio therefore does not track system consolidation in any direction. It is virtually identical in the expansive and the consolidated clusters (1.79 in both), despite their opposite positions on the consolidation gradient. It is slightly lower in the transitional configuration (1.66) and falls below unity in the fragile and intermediate ones (0.75 and 0.70). Figure 6 summarizes these profiles, with the segment between the two markers indicating the direction and size of the imbalance in each cluster.

In response to the third research question, the safest reading is a more circumscribed one. Cluster 2 shows a distinctly primary-heavy accentuation, consistent with its role in absorbing basic-education demand. Across the remaining configurations, the primary-to-

[9] Figure 6: primary–secondary profile by cluster, dumbbell plot

secondary balance is heterogeneous rather than tracking a single gradient of consolidation. It is this interplay of location and composition — rather than a strict, uniform differentiation — that I propose to call differentiated complementarity, understood as an exploratory interpretive label rather than a fully confirmed hypothesis.

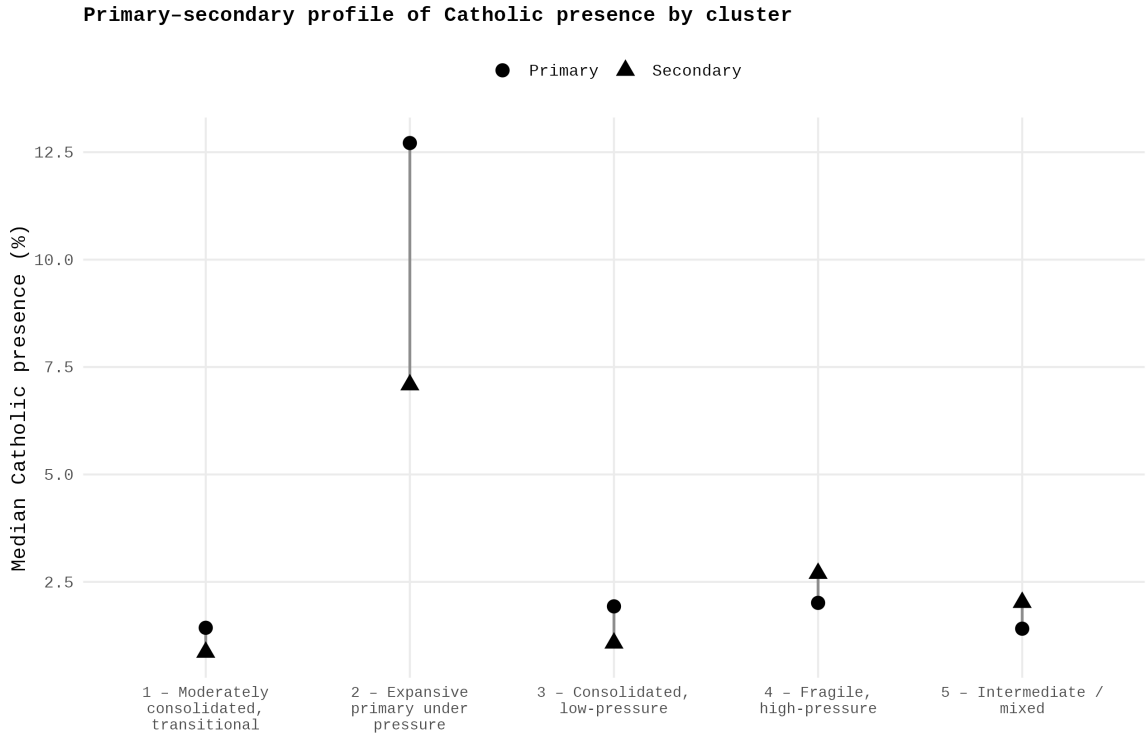


Figure 6. Median Primary–Secondary Profile of Catholic Presence by Cluster (medians computed on non-missing cases; n ranges from 15 to 60 per cluster)

5.5 Absolute versus Relative: Substitution Re-emerges in Qualified Form

It remains to establish whether this pattern reflects system configurations or, more simply, the geography of the Catholic population. The concern is well founded. Cluster 2 also shows the highest median share of Catholic population (31.4%, compared with 11.5% in Cluster 3, 6.2% in Cluster 4, 5.8% in Cluster 1, and 5.6% in Cluster 5), and the Catholic population share is the single strongest predictor of overall educational presence. The historical-demographic explanation thus carries real weight and must be explicitly acknowledged.

A regression of overall presence (log-transformed) on cluster membership, net of the Catholic population share, confirms that the Cluster 2 differential relative to consolidated systems remains large and highly significant once religious demography is held constant. The reported estimate is the one pooled across the five imputed datasets via Rubin's rules ($b = 1.72$, $SE = 0.30$, $p < .001$). None of the variables entering the model is among the imputed indicators, so this estimate coincides exactly with the one obtained on a single completed dataset.

Including the Catholic share raises R^2 from 0.13 to 0.51.[10] Holding religious demography constant, expansive systems under pressure exhibit systematically higher Catholic educational presence. The other three clusters (transitional, fragile, intermediate) do not differ significantly from the consolidated reference net of the Catholic population share, though the fragile cluster approaches conventional significance ($p = .08$). Second, relative penetration — the share of students enrolled in Catholic institutions relative to the share of Catholics in the population — offers a clearer, purely descriptive gradation. It is lowest in the transitional systems of Cluster 1 (0.07 in primary) and in consolidated systems (Cluster 3: 0.13), and it reaches its maximum in the expansive and the fragile clusters alike (0.32 in primary in both). The fragile systems of Cluster 4 stand apart at the secondary level (0.43, against 0.21 in Cluster 2), where the Church educates well beyond its own demographic base. In other words, in absolute terms, Catholic presence peaks in expansive configurations, but relative to its own constituency, it is maximally salient precisely in the most fragile contexts — evidence that recovers, in qualified form, an element of the reading of substitution.

6 Discussion

In this study I have addressed a double invisibility: that of Catholic education in the macro-comparison of education systems, and that of the educational sphere in studies of religion and welfare. By linking, for the first time on a global scale, data from the *Annuarium Statisticum Ecclesiae* to international statistical data sets, I have reconstructed five recurring configurations of national education systems and observed how Catholic educational presence is distributed within them.

The results present a picture more nuanced — and more interesting — than any single rival reading would have predicted in isolation. The dominant pattern in absolute terms is one of institutional complementarity. Catholic presence clusters where public systems are expanding under pressure that exceeds their consolidation capacity, not where needs are most extreme nor where systems are most mature.

The historical-demographic explanation carries real weight — the Catholic population share is the single strongest predictor of presence. However, it does not exhaust the pattern: net of religious demography, the differential of expansive configurations remains large and statistically significant. The reading of substitution, rejected in its mechanical version, re-emerges in qualified form when presence is related to its demographic base. In the most fragile systems, the Church educates well beyond its own constituency, with maximum relative penetration.

Finally, on the modulation of form: H3 is not supported in its original formulation. The primary-to-secondary balance does not shift monotonically with system consolidation, since

[10] 95% CI: 0.40–0.61; full coefficient table with confidence intervals in Table S4, Supporting Information

the consolidated cluster remains markedly primary-heavy while the fragile and intermediate clusters are secondary-leaning. This heterogeneity nonetheless suggests that the form of Catholic presence varies systematically across configurations, a pattern I interpret exploratorily, not as a confirmed hypothesis, under the label of differentiated complementarity, which operates not as an alternative to but in combination with the preceding readings.

For the sociology of religion, this outcome carries a general implication: religious presence in institutional fields is neither a premodern residue destined to recede as the state advances, nor a substitute that flourishes only in the state's absence. Rather, it is an actor whose position — and degree of specialization — appears associated with configurations of public capacity. This pattern is consistent with, though it does not directly test, the principle of subsidiarity. What the data show is a pattern of location, not a direct observation of the underlying mechanism.

This contribution also engages with the literature on religion and welfare in Europe (Bäckström and Davie 2010; Bäckström et al. 2011), extending its logic to a global scale and to the educational sphere. The division of labor between state and church that this literature has documented in European social services finds a systematic counterpart in education, with the added value of an empirical typology that makes it possible to specify in which configurations the ecclesial component assumes which function.

The Church's contribution to education systems becomes visible — and interpretable — only when one analyzes not only where needs are greatest, but how national systems respond to those needs and how ecclesial institutions position themselves within such structural configurations. Rendering religious presence legible within international statistical infrastructures is the condition for restoring to it the place it deserves in the comparative analysis of contemporary institutions.

7 Limitations and Future Directions

This study presents several limitations, each of which points to a corresponding direction for future research.

First, the proposed mechanisms — absorption of excess demand, variation in the primary-to-secondary profile across configurations — are not measured directly. The design is ecological and cross-sectional, and the identified configurations are compatible with historical processes that only longitudinal data can disentangle. The historical series of the Annuarium Statisticum Ecclesiae makes this extension feasible.

Second, the absence of comparable data on other non-state providers — secular nonprofits and those of other denominations — prevents establishing whether the observed pattern is specifically Catholic or common to the non-state sector as a whole. This specificity will need to be verified once adequate data become available.

Third, the analysis does not incorporate the legal-financial regimes of Church-State relations (e.g., public funding of Catholic schools, concordats, regulation), which constitute the most plausible institutional mechanism behind the observed complementarity and whose integration represents the natural next step.

Fourth, two variables (the pupil-teacher ratios) show high shares of missing values (56% and 65%). The sensitivity analysis excluding them (§4.1), however, indicates reasonable stability of the factor structure and the cluster partition. Missingness also differs across the outcome measures. The median Catholic secondary presence rests on between 15 and 60 countries depending on the cluster, with the high-pressure configurations the most affected, so the primary-to-secondary comparisons in §5.4 should be read with that in mind.indicates reasonable stability of the factor structure and the cluster partition. Fifth, the reference year was selected independently for each variable in order to maximize country coverage (§4.1). The specific calendar year underlying a given indicator can therefore differ somewhat from that of other indicators for the same country. I consider this an acceptable trade-off against the alternative of a much smaller, single-year sample. It does mean, however, that the analysis should be read as describing a recent cross-section of these education systems rather than a single, precisely dated year. Besides, the statistics do not vary radically from year to year. Changes occur gradually over time.

Finally, the national unit of analysis conceals subnational heterogeneity — urban/rural, regional — that is likely relevant for an actor as capillary as the Church. Extension to subnational data, where available, constitutes a further natural direction.

These limitations do not undermine this study's main point: rendering religious presence legible within international statistical infrastructures is the condition for restoring to it the place it deserves in the comparative analysis of contemporary institutions.

Conflicts of Interest

The author declares no conflicts of interest.

References

Bäckström, A., and G. Davie, eds., with N. Edgardh and P. Pettersson. 2010. *Welfare and Religion in 21st Century Europe: Volume 1. Configuring the Connections*. Farnham: Ashgate.

Bäckström, A., G. Davie, N. Edgardh, and P. Pettersson, eds. 2011. *Welfare and Religion in 21st Century Europe: Volume 2. Gendered, Religious and Social Change*. Farnham: Ashgate.

Bryk, A. S., V. E. Lee, and P. B. Holland. 1993. *Catholic Schools and the Common Good*. Cambridge, MA: Harvard University Press.

Casanova, José. 1994. *Public Religions in the Modern World*. Chicago: University of Chicago Press.

Casanova, José. 2019. “Global Religious and Secular Dynamics.” *Brill Research Perspectives in Religion and Politics* 3 (1–4).

Cattell, R. B. 1966. "The Scree Test for the Number of Factors." *Multivariate Behavioral Research* 1 (2): 245–276. https://doi.org/10.1207/s15327906mbr0102_10

Chaves, Mark, and William Tsitsos. 2001. “Congregations and Social Services: What They Do, How They Do It, and With Whom.” *Nonprofit and Voluntary Sector Quarterly* 30 (4): 660–683.

Chaves, Mark, Mary Hawkins, Anna Holleman, and Joseph Roso. 2020. “Introducing the Fourth Wave of the National Congregations Study.” *Journal for the Scientific Study of Religion* 59 (4): 646–650.

Cnaan, R. A. 2002. *The Invisible Caring Hand: American Congregations and the Provision of Welfare*. New York: NYU Press.

Coleman, James S., Thomas Hoffer, and Sally Kilgore. 1982. *High School Achievement: Public, Catholic, and Private Schools Compared*. New York: Basic Books.

D'Agostino, T. J., and M. Niaz Asadullah. 2025. “Faith-Based Education and Development: Opportunities, Challenges, and Controversies.” *International Journal of Educational Development* 119: 103440.

D'Agostino, T. J., and P. G. Carozza. 2019. "Extending the Research Orientation and Agenda for International and Comparative Studies in Catholic Education." *International Studies in Catholic Education* 11 (2): 140–158. https://doi.org/10.1080/19422539.2019.1641048

Day Ashley, L., C. Mcloughlin, M. Aslam, J. Engel, J. Wales, S. Rawal, R. Batley, G. Kingdon, S. Nicolai, and P. Rose. 2014. *The Role and Impact of Private Schools in Developing Countries: A Rigorous Review of the Evidence*. London: Department for International Development (DFID).

Davies, S., and L. Quirke. 2007. “The Impact of School Sector on School Organizations: Institutional and Market Logics.” Sociology of Education 80 (1): 66–89.

Dronkers, J., and S. Avram. 2010. "A Cross-National Analysis of the Relations of School Choice and Effectiveness Differences between Private-Dependent and Public Schools." *Educational Research and Evaluation* 16 (2): 151–175. https://doi.org/10.1080/13803611.2010.484977

Dronkers, J., and P. Robert. 2008. "Differences in Scholastic Achievement of Public, Private Government-Dependent, and Private Independent Schools." *Educational Policy* 22 (4): 541–577. https://doi.org/10.1177/0895904807307065

Elfert, M., and C. Ydesen. 2023. *Global Governance of Education: The Historical and Contemporary Entanglements of UNESCO, the OECD and the World Bank*. Educational Governance Research 24. Cham: Springer. https://doi.org/10.1007/978-3-031-40411-5

Finke, Roger, and Rodney Stark. 1992. *The Churching of America, 1776–1990: Winners and Losers in Our Religious Economy*. New Brunswick, NJ: Rutgers University Press.

Freeman, John, and Michael T. Hannan. 1983. "Niche Width and the Dynamics of Organizational Populations." *American Journal of Sociology* 88 (6): 1116–1145.

Gallego, F. A., and R. D. Woodberry. 2010. "Christian Missions and Education in Former African Colonies: How Competition Mattered." Journal of African Economies 19 (3): 294–319.

Gill, A., and E. Lundsgaarde. 2004. "State Welfare Spending and Religiosity: A Cross-National Analysis." *Rationality and Society* 16 (4): 399–436. https://doi.org/10.1177/1043463104046694

Göçmen, İ. 2013. "The Role of Faith-Based Organizations in Social Welfare Systems: A Comparison of France, Germany, Sweden, and the United Kingdom." *Nonprofit and Voluntary Sector Quarterly* 42 (3): 495–516. https://doi.org/10.1177/0899764013482046

Grace, G. 2016. *Faith, Mission and Challenge in Catholic Education*. London and New York: Routledge.

Grace, G., and J. O'Keefe, eds. 2007. *International Handbook of Catholic Education: Challenges for School Systems in the 21st Century*. Dordrecht: Springer.

Grace, G., and Q. Wodon. 2021. "Catholic and Faith-Based Schools in Sub-Saharan Africa: Introduction to the Special Issue." *International Studies in Catholic Education* 13 (2). https://doi.org/10.1080/19422539.2021.2010448

Hannan, Michael T., and John Freeman. 1977. "The Population Ecology of Organizations." *American Journal of Sociology* 82 (5): 929–964.

Hanushek, E. A., and L. Woessmann. 2007. *The Role of Education Quality for Economic Growth*. World Bank Policy Research Working Paper No. 4122. Washington, DC: The World Bank.

James, Estelle. 1993. "Why Do Different Countries Choose a Different Public-Private Mix of Educational Services?" *Journal of Human Resources* 28 (3): 571–592.

Jolliffe, I. T. 2002. *Principal Component Analysis*. 2nd ed. New York: Springer.

Kaufman, L., and P. J. Rousseeuw. 2005. *Finding Groups in Data: An Introduction to Cluster Analysis*. Hoboken, NJ: Wiley.

Meyer, John W., John Boli, George M. Thomas, and Francisco O. Ramirez. 1997. "World Society and the Nation-State." *American Journal of Sociology* 103 (1): 144–181.

OECD. 2024. *Education at a Glance 2024: OECD Indicators*. Paris: OECD Publishing. https://doi.org/10.1787/c00cad36-en

OECD/European Union/JRC. 2008. *Handbook on Constructing Composite Indicators: Methodology and User Guide*. Paris: OECD Publishing.

Pius XI. 1931. Quadragesimo Anno. Encyclical letter.

Pontifical Council for Justice and Peace. 2004. Compendium of the Social Doctrine of the Church. Vatican City: Libreria Editrice Vaticana.

Pritchett, L. 2015. "Creating Education Systems Coherent for Learning Outcomes: Making the Transition from Schooling to Learning." RISE Working Paper 15/005. https://doi.org/10.35489/BSG-RISE-WP_2015/005

Ragin, Charles C. 2000. *Fuzzy-Set Social Science*. Chicago: University of Chicago Press.

Ragin, Charles C. 2008. *Redesigning Social Inquiry: Fuzzy Sets and Beyond*. Chicago: University of Chicago Press.

Sciffer, M. G., et al. 2026. “Socioeconomic and Indigenous School Segregation in Australia: The Role of Institutional Differentiation and Fees.” British Educational Research Journal.

Steiner-Khamsi, G., S. Appius, and A. Nägeli. 2024. "Comparing Two Transfer Spaces over Time and against a Global Script: The Case of School-Autonomy-with-Accountability." *Journal of Education Policy* 40 (3): 355–375. https://doi.org/10.1080/02680939.2024.2394594

Sullivan, J. 2001. *Catholic Education: Distinctive and Inclusive*. Dordrecht: Kluwer Academic Publishers.

UNESCO. 2021. *Global Education Monitoring Report 2021/2: Non-State Actors in Education: Who Chooses? Who Loses?* Paris: UNESCO Publishing.

Verger, Antoni, Clara Fontdevila, and Adrián Zancajo. 2017. *The Privatization of Education: A Political Economy of Global Education Reform*. New York: Teachers College Press.

Wodon, Q. 2014. *Education in Sub-Saharan Africa: Comparing Faith-Based, Private Secular, and Public Schools*. Washington, DC: The World Bank.

Wodon, Q. 2019. "Symposium on Catholic Schools and the Changing Global Landscape for Faith-Based Education: An Introduction." *The Review of Faith & International Affairs* 17 (4): 48–51. https://doi.org/10.1080/15570274.2019.1681785

Woodberry, R. D. 2012. “The Missionary Roots of Liberal Democracy.” American Political Science Review 106 (2): 244–274.

Supporting Information

The supporting material referenced throughout the article is reproduced in the appendix below.

Mapping the Global Catholic Educational Presence: Structural Configurations of National Education Systems

Supplementary Tables and Figures

This appendix reports supplementary technical material referenced throughout the article: the full eigenvalue table for all thirteen principal components (Table S1) and the corresponding scree plot (Figure S1); the unclustered scatterplot of countries in PC1–PC2 space (Figure S2); the complete table of variable loadings on the first two components (Table S2); the fuzzy cluster validity indices for k = 2 to 6 (Table S3); the full pooled regression table with confidence intervals (Table S4); cluster sizes and the distribution of fuzzy membership degrees (Table S5); and the country composition of the five clusters (Table S6).

Table S1. Eigenvalues and Explained Variance (All 13 Components)

Component	Eigenvalue	Variance Explained (%)	Cumulative Variance
PC1	6.225	47.89	47.89
PC2	1.670	12.85	60.73
PC3	1.026	7.89	68.63
PC4	0.939	7.22	75.85
PC5	0.820	6.31	82.16
PC6	0.670	5.15	87.31
PC7	0.497	3.82	91.13
PC8	0.456	3.51	94.64
PC9	0.255	1.96	96.60
PC10	0.157	1.21	97.80
PC11	0.149	1.14	98.95
PC12	0.087	0.67	99.62
PC13	0.049	0.38	100.00

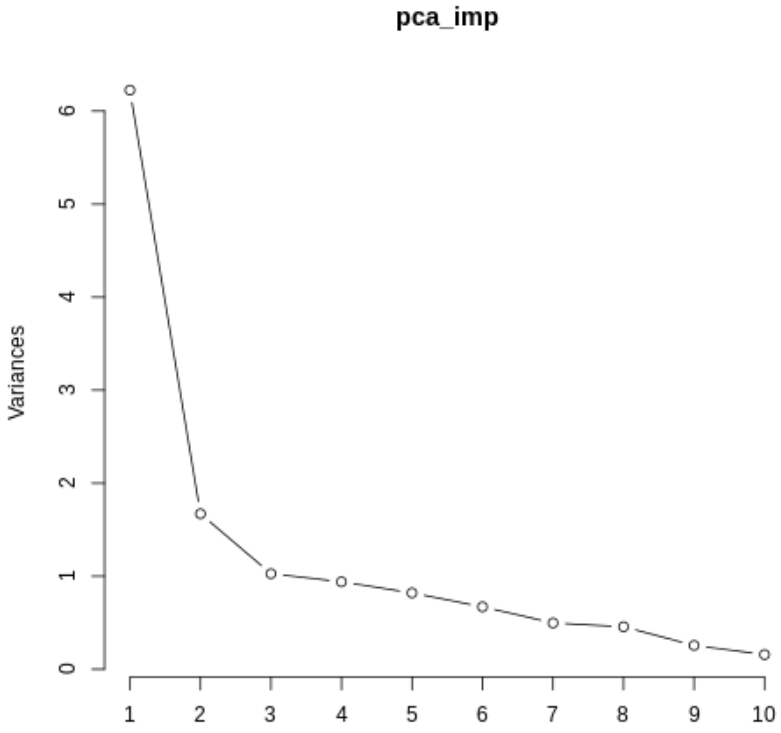


Figure S1. Scree Plot of the 13 Principal Components

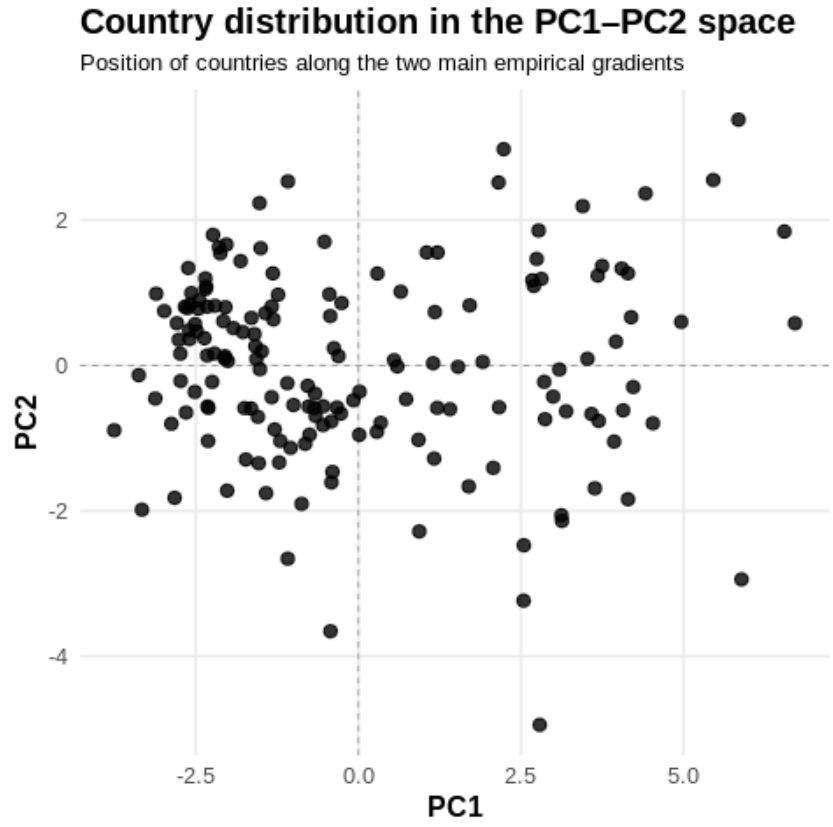


Figure S2. Distribution of Countries in PC1–PC2 Space

Table S2. Eigenvector Coefficients ("Loadings") of the First Two Principal Components (All 13 Variables)

Variable	PC1	PC2
Pupil-teacher ratio, primary	0.365	−0.060
Share of children in population	0.364	−0.142
Infant mortality	0.363	−0.053
Primary students per 1,000 inhabitants	0.322	−0.384
Population growth rate	0.252	−0.001
Pupil-teacher ratio, secondary	0.264	−0.154
Secondary students per 1,000 inhabitants	−0.157	−0.481
Secondary enrollment	−0.364	−0.063
Primary completion rate	−0.316	−0.269
Secondary teachers (% of population)	−0.273	−0.072
Public spending on education (% of GDP)	−0.155	−0.083
Primary teachers (% of population)	−0.102	−0.331
Primary enrollment	−0.012	−0.611

Table S3. Fuzzy Cluster Validity Indices for k = 2 to 6

PC = fuzzy partition coefficient (higher indicates a crisper partition); PE = partition entropy (lower indicates a crisper partition); XB = Xie–Beni index (lower indicates better-separated, more compact clusters); FS = fuzzy silhouette approximation (higher indicates better-separated clusters). Both PC and PE decline/rise monotonically with k, a well-known property of these indices that limits their usefulness for selecting k on their own. The Xie–Beni index is minimized at k = 5 on the two-component solution (PC1–PC2), providing the main quantitative basis, alongside substantive interpretability and stability across initializations, for retaining five clusters; on the three-component solution (PC1–PC3) the Xie–Beni index is close between k = 5 and k = 6, without a clearly dominant solution.

Indices computed on PC1–PC2

k	PC	PE	XB	FS
2	0.822	0.3	0.508	0.554
3	0.697	0.539	0.637	0.41
4	0.651	0.662	0.493	0.398
5	0.596	0.805	0.406	0.394
6	0.545	0.933	0.463	0.345

Indices computed on PC1–PC3

k	PC	PE	XB	FS
2	0.781	0.36	0.683	0.49
3	0.618	0.663	0.954	0.322
4	0.547	0.843	0.753	0.295
5	0.498	0.993	0.549	0.309
6	0.451	1.139	0.541	0.308

Table S4. Regression of Overall Catholic Educational Presence (Log-Transformed) on Cluster Membership and Catholic Population Share, Pooled Across the Five Imputed Datasets (Rubin's Rules)

Reference category: Cluster 3 (consolidated, low-pressure). N = 165. Dependent variable: log(cep_primary_secondary + 0.1). Pooled $R^2 = 0.510$ (95% CI: 0.397–0.611). None of the variables entering the model is among the thirteen indicators subjected to multiple imputation (see §4.1): the between-imputation variance is therefore null, the pooled estimates coincide with those obtained on a single completed dataset, and the degrees of freedom correspond to the complete-data residual degrees of freedom (165 − 6).

Term	Estimate	SE	t	df	p	95% CI low	95% CI high
Intercept	-1.18	0.18	-6.568	159	<.001	-1.535	-0.825
Cluster 1 (transitional)	-0.259	0.259	-0.997	159	.320	-0.771	0.254
Cluster 2 (expansive under pressure)	1.721	0.304	5.662	159	<.001	1.12	2.321
Cluster 4 (fragile, high-pressure)	0.557	0.316	1.761	159	.080	-0.068	1.183
Cluster 5 (intermediate/mixed)	0.321	0.279	1.15	159	.252	-0.23	0.872
Catholic population share (%)	0.032	0.003	11.172	159	<.001	0.027	0.038

Table S5. Cluster Sizes and Fuzzy Membership Distribution

Cluster	Label	N
3	Consolidated, low-pressure	60
1	Moderately consolidated, transitional	35
5	Intermediate / mixed	28
2	Expansive primary under pressure	22
4	Fragile, high-pressure	20

Distribution of maximum membership degree across the 165 countries (fuzzy c-means, k = 5):

Min	1st Qu.	Median	Mean	3rd Qu.	Max
0.34	0.561	0.74	0.717	0.881	0.997

27 countries (16.4%) have a maximum membership below 0.5, and 46 (27.9%) below 0.6, indicating a substantial share of hybrid or transitional cases consistent with the rationale for using fuzzy rather than hard clustering.

Table S6. Country Composition of the Five Clusters (Hard Assignment)

Country lists correspond to hard assignment (maximum membership); see Section 5.2 for cluster labels.

Cluster 3 – Consolidated, low-pressure (n = 60): Malta; Luxembourg; Macao SAR; Cyprus; Mauritius; Estonia; Bahrain; Latvia; Slovenia; North Macedonia; Jamaica; Albania; Lithuania; Armenia; Moldova; Bosnia and Herzegovina; Uruguay; Georgia; Croatia; Kuwait; New Zealand; Slovakia; Norway; Finland; Singapore; Denmark; Bulgaria; Serbia; Hong Kong SAR; Ireland; Switzerland; Austria; Belarus; Hungary; Azerbaijan; Greece; Portugal; United Arab Emirates; Czechia; Cuba; Belgium; Netherlands; Romania; Chile; Australia; Poland; Canada; Ukraine; Spain; Korea (Rep. of); Italy; France; Great Britain; Thailand; Germany; Türkiye; Japan; Russian Federation; United States; China, Mainland

Cluster 1 – Moderately consolidated, transitional (n = 35): Maldives; Cabo Verde; Suriname; Guyana; Bhutan; Fiji; Mongolia; Oman; Costa Rica; Turkmenistan; Kyrgyzstan; Libya; Israel; Sweden; Dominican Republic; Tunisia; Bolivia; Ecuador; Kazakhstan; Sri Lanka; Nepal; Saudi Arabia; Malaysia; Peru; Venezuela; Uzbekistan; Morocco; Algeria; Argentina; Colombia; Viet Nam; Philippines; Mexico; Brazil; Indonesia

Cluster 5 – Intermediate / mixed (n = 28): Solomon Islands; Comoros; Trinidad and Tobago; Timor-Leste; Gabon; Gambia; Botswana; Qatar; Lebanon; Panama; El Salvador; Nicaragua; Lao PDR; Paraguay; Honduras; Tajikistan; Jordan; Cambodia; Guatemala; Syrian Arab Republic; Ghana; Iraq; Myanmar; South Africa; Iran; Egypt; Bangladesh; India

Cluster 2 – Expansive primary under pressure (n = 22): Eswatini; Guinea-Bissau; Lesotho; Mauritania; Togo; Papua New Guinea; Sierra Leone; Haiti; Benin; Burundi; Rwanda; Zimbabwe; Zambia; Malawi; Cameroon; Côte d'Ivoire; Madagascar; Mozambique; Afghanistan; Uganda; Kenya; Congo (Dem. Rep. of the)

Cluster 4 – Fragile, high-pressure (n = 20): Djibouti; Equatorial Guinea; Eritrea; Liberia; Central African Republic; Congo; Guinea; South Sudan; Chad; Senegal; Mali; Burkina Faso; Niger; Yemen; Angola; Sudan; Tanzania; Ethiopia; Nigeria; Pakistan