DOI: 10.1002/epa2.1172

ORIGINAL ARTICLE

EPA EUROPEAN POLICY ANALYSIS

Advancing the operationalization of national policy styles

Nikolaos Zahariadis[1] | **Vassilis Karokis-Mavrikos**[2] | **Theofanis Exadaktylos**[2] | **Alexandros Kyriakidis**[3] | **Jörgen Sparf**[4,5] | **Evangelia Petridou**[4,5]

[1]Rhodes College, Memphis, Tennessee, USA

[2]Department of Politics, University of Surrey, Guildford, Surrey, UK

[3]Center for Research on Democracy and Law, University of Macedonia, Thessaloniki, Greece

[4]Risk and Crisis Research Center, Mid Sweden University, Östersund, Sweden

[5]NTNU Social Research, Trondheim, Norway

Correspondence
Evangelia Petridou, NTNU Social Research, Thessaloniki, Norway.
Email: evangelia.petridou@miun.se

Abstract

While national policy styles have (re)gained academic attention in recent comparative public policy work, the concept still needs a widely accepted operationalization that can allow the collection and analysis of data across contexts while steering away from construct validity threats. We build on Tosun and Howlett's (2022) work and employ a mixed-methods approach, which relies on exploratory factor analysis and hierarchical cluster analysis. We put forth an operationalization, using Bertelsmann's Sustainable Governance Indicators (SGI) as proxies, that achieves conceptual clarity and distinctiveness, informational robustness, and statistical power. Ultimately, we construct two composite indicators—mode of problem-solving and inclusiveness—calculate them in 41 countries and present policy style classifications based on their combinations. We report the distribution of countries across four policy styles (administrative, managerial, accommodative, adversarial) and conclude with an analysis of the clusters, assessments of robustness, and comparison with other national policy style classification schemes.

KEYWORDS
comparative policy analysis, hierarchical cluster analysis, Policy styles, quantitative policy analysis, sustainable governance indicators

1 | INTRODUCTION

While the concept of national policy style has (re)gained academic attention in recent comparative public policy work (Howlett & Tosun, 2019; Leiren et al., 2021; Oztig, 2022; Zahariadis et al., 2021, 2022), analytical progress has been hampered by the lack of systematic applications across a large number of countries. This may be partly due to the inability to quantitatively operationalize the concept in ways that permit the collection and analysis of data in different national contexts. Tosun and Howlett (2022) aim to change that by utilizing indicators taken from Bertelsmann's Sustainable Governance Indicators (SGI) project. They argue convincingly that some of the indicators found in the 'good governance' grouping of that database, specifically strategic planning and public consultation, are useful proxies for measuring the two dimensions of policy styles: mode of problem-solving and state-society relations. While their argument is theoretically and empirically beneficial, especially in terms of steering the quantitative literature forward, their recommendation regarding operationalization is problematic due to construct validity threats. We argue that the indicators treated in Tosun and Howlett (2022) fall short of operationalizing and measuring the two dimensions of policy styles and we instead propose a new set of composite variables to quantitatively operationalize and measure national policy styles. We are not the first to propose using SGI data for this purpose, but instead, we advance the operationalization of national policy styles. Although we retain the core of Tosun and Howlett's (2022) operationalization argument and the conceptual value of SGI's good governance measures as proxy indicators (referred to simply as SGI hereinafter), we chart a different direction that overcomes the limitations posed by construct validity threats, carefully calibrating the effects of three criteria: conceptual clarity and distinctiveness, informational robustness, and statistical power. After reviewing different conceptualizations of policy styles and their limitations, we operationalize the concept and examine how our new measurement affects classificatory expectations of national policy styles. Applying exploratory factor analysis (EFA) and hierarchical cluster analysis (HCA) to the SGI data, we amend Tosun and Howlett's (2022) argument to demonstrate that more SGI data should be fruitfully used to operationalize and measure policy styles as tools of long-term patterns of policy-making, including comparative analyses.

2 | DIFFERENT CONCEPTUALIZATIONS AND MEASUREMENTS OF NATIONAL POLICY STYLES

The concept of a national policy style focuses on the institutional arrangements that govern policy deliberations and how policy actors interact with each other within a broader ideational framework. These interactions provide actors with a blueprint through which policy-making is filtered in terms of path dependencies (Pierson, 2004). Hence, the modes of deliberation—policy styles—affect policy-making alongside previous decisions. Tackling the concept of a

policy style Richardson (1982) suggested two dimensions in these interactions: (a) the way governments approach problem-solving, either anticipating or reacting to problems, and (b) the way government and society interact, either through consensus-building or more top-down decisions. Out of this two-by-two matrix conceptualization, four policy styles emerge, under which countries could be classified. The main issue, however, was that Richardson's original framework could not capture the complexities of national political systems, especially when it came to the role of bureaucrats and experts vis-à-vis the public, as well as the degrees of inclusion of such actors within decision-making. At the same time, it was restricted to European or, more broadly speaking, western liberal democratic political systems. Howlett and Tosun's (2019) elaboration focuses precisely on these criticisms, accepting Richardson's premise and proposing a measurement that looks at key policy actors as a variable expressed as high/low inclusiveness of societal actors in policy-making. Doing so enables the authors to apply the concept beyond Europe to major developing and developed countries.

In a recent study of the interaction of policy styles with political trust, Zahariadis et al. (2022) revisit Richardson's original argument and use Howlett and Tosun's reconceptualization as a springboard to expand and refine the concept to measure the capacity of governments to solve problems. They focus on how Richardson's "ways of doing things" idea (Mazey & Richardson, 1993) manifests in governments' policy-making capacity. For instance, if the style is anticipatory in terms of problem-solving, it is reasonable to expect some sort of monitoring or implementation capacity. In other words, beyond the actual institutional focus, this must also be carefully circumscribed. Zahariadis et al. (2022) retain the two dimensions of policy style but propose administrative arrangements as the variable of interest (high/low policy capacity) and state-society relations (high/low inclusiveness). Policy-making is strongly affected by existing administrative arrangements (more than bureaucratic structures, which remain largely unchanged) and determines a government's approach to problem-solving, shaping strategic opportunities and available alternatives. But their argument remains qualitative in measurement; the authors do not propose any quantitative scores for policy style.

Tosun and Howlett's (2022) quantitative operationalization of the concept of policy style is a major step forward. However, measuring "ways of doing things" exclusively as planning capacity does not capture the construct's richness. While the state apparatus may include strong strategic planning capabilities, these may not be useful or productive in the policy process because of a lack of administrative capabilities, such as the inability to implement decisions via the existing administrative apparatus. Richardson, Gustaffson, et al. (1982, p. 2) categorically assert that policy style "also encompasses policy execution."

State-society relations, as per Howlett and Tosun (2019), refer to accommodative—or not—patterns of interest mediation. Tosun and Howlett (2022) measure these patterns as social consultation. What matters to Zahariadis et al. (2022) goes back to Richardson (1982). The authors argue that patterns of mediation refer to the degree of inclusiveness of social actors in policy-making and how institutionally formalized this degree is. On the one hand, in cases where social actors are more meaningfully included in the policy process, consensus—or, at least, mutual understanding—is built in the policy formulation and adoption stages, aiming at reducing the potential for conflict at the implementation stage. On the other hand, in political systems where minimal social consultation takes place, decisions are usually directed from the top down. In addition to the level of consultation, there also exist different modes of societal consultation. For example, consultation manifests not only in institutionalized arrangements such as formal state-society hearings and remittances but also in the system's transparency, and the quality of information citizens receive that enables such participation (e.g., Balla, 2015).

Nuances within each system that are dictated by previous administrative traditions, or the historical evolution of political relations, affect the level and mode of consultation. Therefore, the degree of inclusiveness, which captures the broad range of citizen and interest group interaction with government in the policy-making and implementation processes, is the optimal way to operationalize state-society relations and the one closest to Richardson's original conceptualization.

Policy style necessitates a broader, more complete quantitative operationalization, which is feasible using SGI data. We, therefore, advance the operationalization argument by Tosun and Howlett (2022), suggesting different indicators to convey conceptual clarity to capture the informational robustness of the concept of policy style and strengthen its statistical power.

3 | OPERATIONALIZING POLICY STYLE WITH SGI INDICATORS

Taking a step back from Tosun and Howlett's (2022) work, we revisit the question of how best to numerically operationalize the concept of national policy style and overcome construct validity threats. Our analysis rests on the premise that the SGI indicators—we concentrate only on those, as Tosun and Howlett (2022) do—contain valuable information pertinent to national policy styles. Although SGI indicators have a solid theoretical basis and incorporate, for our purposes, important numerical data measuring the government's "capacity to steer [policy] processes with success," they also have limited temporal and geographical coverage in addition to problems with constructing aggregate indicators from individual attributes among others (Schraad- Tischler & Seelkopf quoted in Croissant & Pelke, 2022, p. 144). While realizing they are proxy indicators, we also acknowledge they are imperfect measures because the data were not collected for this purpose.

To answer the question of which indicators are best suited to the operationalization of policy style, we need to address two major issues: construct validity and methodological problems that afflict composite variables. Construct validity is an important threat that has the potential to plague every research design: does the indicator measure the concept it says it is measuring (Cronbach & Meehl, 1955)? For example, a study using GDP as an indicator of national power potentially suffers from construct validity threats because power entails much more than economic output. Similarly, the concept of national policy style is more than strategic planning at the highest government level. A "way of doing things" also involves the execution of plans and the capacity to do so (Richardson, 1982).

Composite variables are additionally vulnerable to methodological problems. We use the term composite variable to refer to the mathematical combination of individual component indicators (Nardo et al., 2005, p. 7). Such variables may "send misleading or non-robust policy messages if they are poorly constructed or misinterpreted" (Nardo et al., 2005, p. 6). For this reason, it is essential to use a measurement process that ensures quality and transparency. We follow a framework proposed by Freudenberg (2003), which consists of five stages: theory, selection of variables, standardization, weighting, and robustness.

To address construct validity threats and measurement issues, we collapse the above stages into three phases/criteria: conceptual distinctiveness and clarity, informational robustness, and statistical power. The first addresses the most critical stage, theory. As Freudenberg (2003, p. 7) asserts: "A theoretical framework is needed to combine individual indicators into a meaningful composite and to provide a basis for the selection of components and weights." Informational

robustness encapsulates the selection of variables. After all, composite indicators are summary measures of different dimensions of more abstract constructs which are not directly measurable. Selecting meaningful components is key to constructing a solid composite variable. Statistical power incorporates the remaining stages. Standardization refers to normalizing component indicators measured in different units and scales, which is not a problem here. We focus on weighting and robustness of measurement by conducting tests to ensure data validity and reliability.

3.1 | Conceptual distinctiveness and clarity

Our theoretical departure point is the concept of national policy styles by Richardson (1982) as enriched by refinements by Howlett and Tosun (2019) and Zahariadis et al., (2021, 2022). Our task is to quantitatively operationalize the approach or mode of problem-solving and inclusiveness. Mode of problem-solving is one of the two conceptual dimensions proposed by Richardson to capture the government's "way of doing things." It summarizes patterns of interaction among policy actors within specific national institutional frameworks. While useful, the concept of policy capacity by Zahariadis et al. (2022) may be a bit more institutional in nature because it refers to the prerequisites (competencies and capabilities) needed in the policy-making process (Wu et al., 2015, p. 165). This entails that competency (legal, administrative, and human) is in place and coordinated (at the right place at the right time) so that the government may perform policy functions effectively (Peters, 2015).

For this reason, we go back to Richardson's original policy process conceptualization and examine its policy-making (planning and adoption), and policy implementation dimensions. Unlike Richardson, we offer a numerical, instead of a dichotomous, measure of mode and a more nuanced conceptualization because we explicitly measure the three components of the construct. We measure the capacity of governments to think through problems and plan courses of action; the ministerial politics of adopting particular solutions to these problems; and of course the administrative capacity to execute them.

Conversely, as a proxy for state-society relations, inclusiveness indicates the degree to which societal actors have a seat at the policy-making table and the degree of formalization of their participation (Zahariadis et al., 2022). These are conceptually distinct dimensions of policy style: *whereas mode of problem solving is a feature of a state's political system and bureaucratic apparatus, inclusiveness foregrounds that system's ability to meaningfully and formally incorporate the input of its citizens in the decision-making processes.* To operationalize and measure policy styles cross-nationally, we turn to SGI, as Tosun and Howlett (2022) did, but we select different numerical indicators so that we may address construct validity issues.

The SGI project is based on the normative notion that governments must make policies that address the challenges they face—globalization adversities, social cleavages, resource limitations, and demographic shifts—while considering the long-term effects of the decisions they make (Bertelsmann Stiftung, 2022). The project aims to inform policy-making in 41 OECD countries through cross-national comparisons based on a customized set of indicators forming three main pillars (or indices): policy performance, robust democracy, and good governance (Schraad-Tischer & Seelkopf, 2015). In our study, we use indicators from the good governance pillar, guided by the broad question, "[h]ow well developed are the governance capacities of OECD/EU countries in terms of interaction between government and societal actors?" (Schiller et al., 2022, p. 7), which is also at the core of the policy styles concept. SGI conceptualizes

governance as both the capacity of a government to act as well as how institutions and societal actors influence what the government does. The good governance pillar is divided into two dimensions: *executive capacity* and *executive accountability*, under which various indicators exist; we focus on three indicators to measure the mode of problem-solving and an additional three to measure inclusiveness at the policy process level.

More specifically, we operationalize the mode of problem-solving with the numerical indicators of strategic capacity, interministerial coordination, and policy implementation. They correspond to the policy process conceptualization of planning, the politics of adoption, and the delivery of government services. To operationalize inclusiveness, we use societal consultation, policy communication, and citizens' participatory competence. These indicators, except societal consultation and policy communication, are, *per se*, composite indicators. Here we break with Tosun and Howlett (2022); we employ the entire composite indicator in our analysis rather than selecting some of its components. We recognize that a potential SGI drawback is the occasional aggregation of indicators to higher-level concepts without adequate theoretical justification (Croissant & Pelke, 2022, p. 155). Yet, we still reasonably argue that the components of each composite indicator were designed to measure as many of its facets as possible.

3.2 | Informational robustness

3.2.1 | Mode of problem solving

We argue that the mode of problem-solving may be measured using *strategic capacity, interministerial coordination*, and *implementation*. The three components cover in large part three important elements of the policy process: planning, adoption, and implementation. Strategic capacity consists of two components: strategic planning and expert advice. Strategic planning concerns the role institutionalized forms of strategic planning play in the long-term view of policy challenges. This goes hand-in-hand with the second component, expert advice, which measures the degree to which these institutions regularly consider the advice of experts in their decision-making processes.

Interministerial coordination addresses the politics of negotiating line ministry proposals against the priorities the government has set. The components of this indicator include the level of expertise of the government offices; the involvement of line ministries; the ability of cabinet committees to filter out and settle issues before cabinet meetings; the extent that ministerial officials and civil servants have the ability to coordinate the drafting of policy proposals with other ministries before proposals reach ministerial coordination; the existence of information coordination mechanisms, and the level of digitalization for interministerial coordination.

In turn, interministerial coordination concerns political coordination that links planning and execution. Implementation refers to the effective execution of government policies. Its components measure the extent to which the government may achieve its policy objectives; the degree to which there are mechanisms to ensure ministerial compliance; the degree to which there are instruments to monitor ministries and public agencies; the degree to which the national government ensures that the tasks allocated to subnational governments are adequately funded; the degree of decentralization of subnational units of government; the degree to which the national government ensures that subnational governments realize national priorities of public services, and finally, the ability to deal with powerful lobby groups when it comes to lobby groups' influence in the policy-making process.

Using the composite indicator allows us to nuance the concept of the mode of problem-solving. It overlaps substantially with policy capacity (Zahariadis et al., 2022) but is closer to Richardson's conventional policy process conceptualization. We parse it to different levels of government and their attendant bureaucracies, as well as capture the relationships among them—both formal and informal—thus boosting informational robustness, which in turn increases construct validity. The indicators and their components measure an exhaustive number of aspects of policy problem-solving from different perspectives of the political system.

3.2.2 | Inclusiveness

We capture inclusiveness through *societal consultation, policy communication*, and *citizens' participatory competence*. These are important components of policy style measuring the interpenetration between state and societal actors in the policy process. Societal consultation assesses how successful the government is in inviting and considering the opinions of societal organizations such as trade unions, business associations, religious communities, and others. Policy communication addresses whether the government communicates its policy goals in a factually coherent way. To complete the operationalization of inclusiveness, we include an indicator that aims to assess the extent to which citizens are well informed of public policies and the extent to which the government publishes data and information in such a way that it strengthens the ability of citizens to hold the government accountable. Even though we do not measure accountability, this indicator strengthens the operationalization and measurement of inclusiveness because it aims to assess how well the citizenry is informed, which in turn can increase the level and quality of inclusiveness.

4 | TWO COMPOSITE VARIABLES MEASURING NATIONAL POLICY STYLE

While these groups of indicators make theoretical sense, how should they be constructed, and are they also statistically valid and reliable?

4.1 | Constructing the composite variables

Mode of problem solving is a composite variable, which is a linear aggregation of the SGI indicators of strategic capacity, interministerial coordination, and implementation. Inclusiveness is a composite variable, which is a linear aggregation of the SGI indicators of societal consultation, citizen participatory competence, and policy communication. We acknowledge that the theoretical framework on policy styles is not sufficiently developed to assign weights to the selected components indicators. For this reason, each composite variable's value is the mean of its component indicators. Table 1 provides the descriptive statistics of the composite variables as well as correlations among the component indicators.

Beyond the usual descriptive statistics of centrality and dispersion, we also report the shape (normality) of our data distribution (Table 1). Skewness and kurtosis are important because, at certain values, they violate the assumption of many frequently used statistics, such as *t*-tests or ordinary least squares regression, leading to a high incidence of Type I errors (Cain et al., 2017).

TABLE 1 Descriptive statistics of composite variables and correlations among their components.

Variable	Mean	Standard deviation	Min	Max	Skewness	Kurtosis	Cronbach's α
Mode of problem solving	6.25	1.23	3.60	8.50	−0.26	2.34	0.86
Inclusiveness	6.20	1.49	3.67	9.17	0.06	2.23	0.85

Correlation matrix

Variable	SCA	IMC	SOC	POL	IMP	CIT
Strategic capacity (SCA)	1.00	-	-	-	-	-
Interministerial coordination (IMC)	0.66	1.00	-	-	-	-
Societal consultation (SOC)	0.80	0.50	1.00	-	-	-
Policy communication (POL)	0.58	0.54	0.66	1.00	-	-
Implementation (IMP)	0.68	0.69	0.73	0.69	1.00	-
Citizen participation (CIT)	0.73	0.52	0.81	0.51	0.68	1.00

The Shapiro–Wilk test for departure from normality yields a W statistic of at least 0.968, which is not significant. In addition, the joint skewness and kurtosis test for normality yields nonsignificant χ^2 coefficients of 1.40 and 1.49, respectively. The data are normally distributed.

We still need to look for differences across time to address one last point: Do the composite variable values change significantly over the medium term? If they do not, as the theory predicts and Tosun and Howlett (2022) visually verify, our argument gains analytical traction. To supplement visual inspection, a statistical assessment is necessary.

We examine all six component indicators (three for mode of problem solving and three for inclusiveness) and divide the sample into three year groups: 2014, 2017, and 2020. While we know there is a glacial change from year to year in some countries, we assess whether it is statistically significant over the short- to medium-term across our sample—3 and 6 years, respectively. We first run a MANOVA to test for equality of means among the groups, assuming homogeneity. Statistically insignificant coefficients indicate no substantial variance across temporal groups. We then run tests assuming heterogeneity in our data. Again, Wald's χ^2 and likelihood-ratio χ^2 are not significant. We conclude that the results of Tosun and Howlett's (2022) visual inspection of two indicators can be generalized to the broader group of our component indicators over the medium-term, as they find, but also in the short-term. There may be a change in some countries over time, but the indicators are, on average, temporally stable without any statistically significant change.

4.2 | Data validity and reliability issues

We can now address statistical issues of construct validity and composite variable creation. First, we correlate the component indicators with other variables to which they should be theoretically related to address convergent validity. Second, we conduct EFA with the component indicators to ascertain their loadings onto latent factors. If they load onto the same factor, each of which represents the mode of problem-solving or inclusiveness, we have gained

additional credibility for the unidimensionality of our conceptual groupings. Finally, we calculate the scale reliability coefficient, Cronbach's α, to assess the internal consistency of measurement.

4.2.1 | Convergent validity tests

Researchers typically establish construct validity by correlating an indicator of a construct that should theoretically be associated with other measures of the same construct (convergent validity) and a number of other indicators that should theoretically vary independently of it (discriminant validity) (Campbell & Fiske, 1959). Unfortunately, our proxy indicators cannot possibly test for discriminant validity because SGI data were not collected for this purpose. However, we can test for convergent validity because there are several conceptually related indicators. We explore convergent validity by correlating our composite variables, mode of problem solving and inclusiveness, with other indicators in SGI to which they should theoretically be related. Two such indicators are organizational reform and the use of evidence-based instruments. It is important to note that these items should be related but not be identical or different measures of the construct. They should simply correlate strongly with our variables so that we may gain further confidence in having overcome convergent validity obstacles.

The correlations indicate strong relationships between our composite variables and theoretically related variables. Mode of problem solving is positively related to organizational reform ($r = 0.717$). Countries with higher planning, political, and implementation capacity tend to also be more willing to change their structure of governance. Their governments appear able to do so because they regularly monitor performance and are willing to change the institutional arrangements of governance if appropriate.

Inclusiveness relates equally strongly with the use of evidence-based instruments ($r = 0.729$). Governments that are willing to consult with the public more frequently and seek to provide more information and encourage greater participation by citizens also tend to be more likely to use regulatory impact assessment exercises (RIA). It is logical that such a relationship exists because RIAs frequently involve consultation and/or collaboration among stakeholders in a transparent way to increase "the quality of RIAs by providing empirical information about the needs and likely reaction of individuals with regard to a regulatory change" (SGI, 2022). Public consultation and the use of evidence-based instruments seem to go hand-in-hand.

4.2.2 | Exploratory factor analysis and dimensionality

EFA is well suited to statistically answer the question of dimensionality. It is exploratory in the sense that we have no preconceived notions of what the latent factors will be (Gorsuch, 2015). EFA involves five phases: suitability, extraction, retention, rotation, and interpretation.

We first examine the suitability of EFA on the data because, as Watkins (2018, p. 226) puts it, "it is [...] important to verify that the measured variables are sufficiently intercorrelated to justify factor analysis." We do this in two ways. First, we present the correlation matrix across all six proposed component indicators (Table 1). A brief inspection shows sufficient intercorrelation of at least 0.50. Second, Bartlett's sphericity test tells us whether the matrix generated by the factors is an identity matrix, that is, the off-diagonal correlation coefficients are randomly different from zero. If the test coefficient comes back as significant, then we may

proceed with EFA. In our case, the χ^2 of 174.67 with 15 degrees of freedom exceeds statistical significance at the 0.000 level. We also test for sampling adequacy using the Kaiser–Meyer–Olkin Measure. It tells us whether the number of observations is statistically enough for EFA. It should be over 0.50 (Beavers et al., 2013, p. 4); ours is 0.833. We conclude that we may safely proceed with EFA.

Extraction identifies the number of factors and their loadings. Two of the most widely used methods in the social sciences are principal axis factor extraction and principal components factor analysis (PCF) (Fabrigar & Wegener, 2012). PCF is more appropriate in our case because its goal is to summarize the information available from a given set of variables and reduce it to a few components (dimensions) that retain as much of the original variance as possible (Conway & Huffcutt, 2003, pp. 150–151). Finding more than one component raises concerns for our argument of unidimensionality among components in the mode of problem-solving or inclusiveness.

The results show that everything loads onto a single factor/dimension (Table 2). This factor explains 78 percent of the variance when it comes to the mode of problem-solving data (eigenvalue of 2.350) and almost 78 percent in reference to inclusiveness (eigenvalue of 2.329). The scree test, which graphs eigenvalues, shows that only one factor should be retained in each situation. We can proceed with rotation to make the results more interpretable. Rotation is divided into two groups of methods, orthogonal and oblique (varimax and promax, respectively). The first assumes uncorrelated components while the second relaxes this assumption. Given the exploratory nature of our endeavor, we decided to run both rotation methods. The results were identical, so, we report promax. The factor loads are strong on the single component in each situation with coefficients of over 0.80 across all cases, which is much higher than the customary level of 0.50 (Table 2).

In short, we interpret our EFA findings to mean that the data can be reduced statistically to a single dimension. Each composite variable that we propose measures the same more-or-less construct even though individual component indicators within each composite variable may do so in conceptually different ways.

4.2.3 | Reliability of measurement

Finally, we conduct a reliability test to ensure the items contain internal consistency of measurement. Reliability is related to measurement error such that the more internally

TABLE 2 Factor loadings using principal components factor extraction and promax rotation.

SGI indicators	Mode of problem solving	Inclusiveness	Uniqueness
Strategic capacity	0.880		0.225
Interministerial coordination	0.881		0.223
Implementation	0.893		0.203
Citizen participation		0.885	0.108
Societal consultation		0.944	0.217
Policy communication		0.809	0.346
Eigenvalue	2.350	2.329	
Cumulative variance explained	0.783	0.776	

consistent the measures are of a construct, the lower the measurement error will likely be. We use Cronbach's α to make this determination. Each factor, mode of problem solving, and inclusiveness consists of three indicators that measure more-or-less the same thing but in conceptually different ways. Cronbach's α assesses whether there is statistical support for this claim.

The reliability coefficients in both cases reveal high levels of internal consistency, above the widely expected level of 0.70. In the case of the mode of problem solving, the data are highly internally consistently measured. The reliability of measurement stands at 0.86 for the mode of problem solving and 0.85 for inclusiveness (Table 1). The higher α coefficient indicates the items share strong covariance and likely measure the same underlying concept.

5 | REVISITING CLASSIFICATION SCHEMES OF NATIONAL POLICY STYLES

All three major theoretical conceptualizations of policy styles (Howlett & Tosun, 2019; Richardson, 1982; Zahariadis et al., 2022) in addition to Tosun and Howlett (2022) classify countries, creating expectations of what their national policy style should be. It is an important first step to generating hypotheses about long-term policy outputs. Having identified robust composite variables of the mode of problem solving and inclusiveness, we want to compare our findings against others to identify similarities and differences. We first examine the data empirically using HCA and then compare the generated clusters of countries to those derived factorially and to other classification schemes.

5.1 | Classification using hierarchical cluster analysis

HCA partitions the data set into groups of units (not factors) such that there is high intra-cluster similarity and high intercluster dissimilarity. It does not tell us why they cluster, just that they do (Everitt et al., 2011, p. 13). More weight is added to the robustness of our argument if clusters identified empirically by HCA significantly overlap with groupings identified by our theory above.

We use a variant of HCA labeled agglomerative analysis whereby clusters are discovered by successive fusions of individuals into groups. Two methods of agglomeration have been found to be most efficient or least problematic: Ward's linkage for similarly sized clusters and average linkage for unequally sized clusters (Ferreira & Hitchcock, 2009). Because HCA is exploratory here and no one method may generally be recommended over others (Gordon, 1999), we decided to run both methods but present only the results from Ward's linkage, which are similar to those obtained via the average linkage. HCA results may be graphically represented in a dendrogram. Read from bottom to top, it reveals several clusters. The question is where to "cut" the tree to capture the most meaningful classification of the data. Usually, this means "a partition such that clusters below that height are distant from each other by at least that amount" although theory and perhaps some subjective criteria may likely be used (Everitt et al., 2011, p. 96).

Figure 1 presents the dendrogram using Ward's linkage. It reveals at least three meaningful groups of countries: Group C, which includes mainly countries of Eastern and Mediterranean Europe (Bulgaria, Croatia, Cyprus, Hungary, Poland, Romania, Slovakia, Greece, Turkey, and

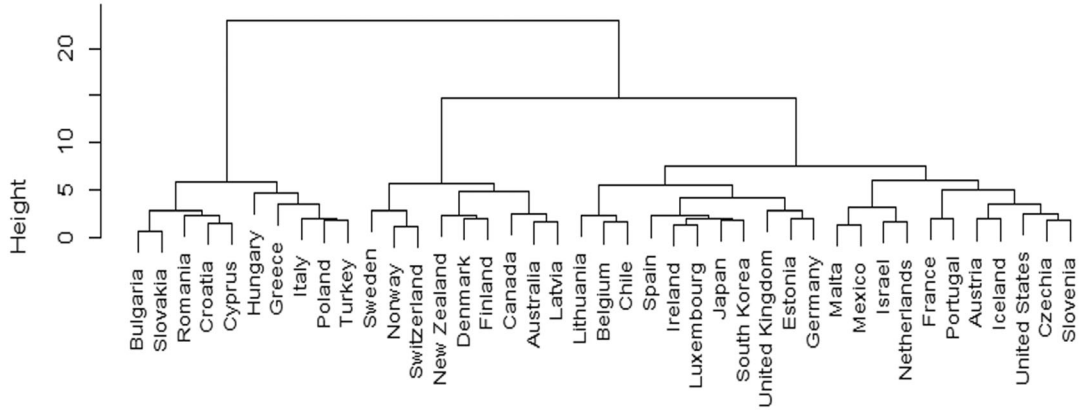

FIGURE 1 Hierarchical Cluster Analysis Dendrogram—Ward's linkage.

Italy), Group B, which includes mainly Nordic countries (Sweden, Norway, Finland, Denmark, Latvia, plus Australia, Switzerland, Canada, and New Zealand), and Group A, which is the rest of the sample. Ward's linkage shows that Group C is more different than the other two, which are more related to each other. We could conceivably explore sub-clusters but because this is mainly a data-driven analysis, we prefer to restrict our findings to broad groups and let theory guide further expectations.

The summary statistics for each group show that the mean across all indicators for Group B is 7.96, which is higher than that of the other two groups, 6.33 for Group A and 4.44 for Group C. The standard deviation of Group B, 0.45, suggests that we have a well-defined and tightly clustered group. Group C appears equally well-defined and clustered with a standard deviation of 0.36. We conclude that our data show a classification scheme that contains at least two well-defined groups at opposite ends of the spectrum. The more voluminous group A, which is more heterogeneous, is situated in between the two groups, with some countries possibly being closer to Group B and a few others closer to Group C. It is important to note this is a two-dimensional space that, for our purposes, delineates the edges but does not give us much information about Group A.

5.2 | Classifying countries according to national policy styles

We are now ready to fit country coordinates in a two-by-two matrix, using the two composite variables for policy style that we derived theoretically (and empirically) above: mode of problem solving and inclusiveness (Figure 2). We compute the averages for both variables: 6.25 for mode and 6.20 for inclusiveness (Table 1) and define as high and low those countries scoring above and below the mean, respectively. The results show a significant overlap between the matrix and HCA findings. Groups B and C in HCA (shown in Figure 2 in bolded italics) land in the expected places. Group B is situated in its entirety at the higher end of the spectrum, as anticipated by HCA (high capacity for problem-solving and high inclusiveness). Group C appears in its entirety at the lower end of the spectrum (low capacity for problem-solving and low inclusiveness). The differences are Slovenia Czechia, Portugal, and Iceland, as well as Austria and Malta—which barely miss the cutoff mean—and are now part of that group. The

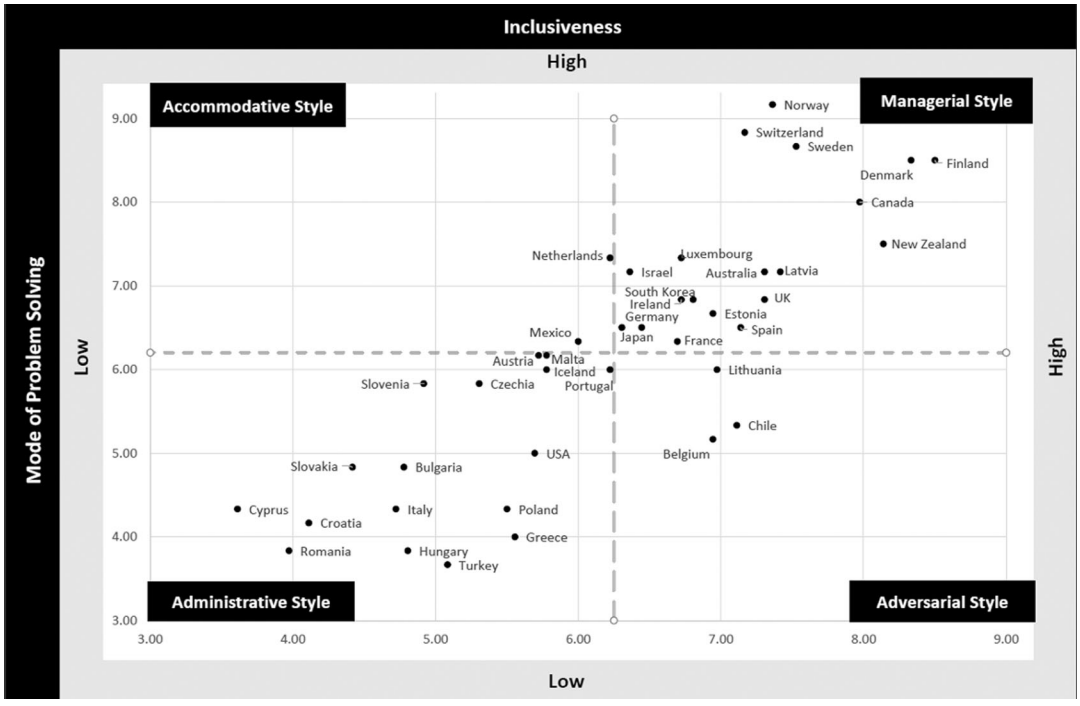

FIGURE 2 Policy styles classification based on inclusiveness and mode of problem-solving scores.

remainder of Group A countries score in-between, with some being closer to Group C and others closer to Group B, as expected. The matrix clarifies and refines this point. The majority of Group A (10 countries) aligns with Group B at the higher end of the spectrum, with some interesting high-low combinations among the remainder of the countries. The two approaches yield roughly similar results, enhancing our confidence in the findings.

We can now revisit the classificatory expectations from other studies to identify similarities and differences. Because of significant conceptual overlap and to facilitate comparisons, we follow Zahariadis et al. (2021) and use the term *administrative style* to refer to a low capacity for problem-solving and low inclusiveness, *managerial style* to refer to high capacity and high inclusiveness, *adversarial style* to refer to high capacity but low inclusiveness, and *accommodative style* to refer to low capacity but high inclusiveness. We first note that the matrix shows only five mixed style cases—three adversarial and two accommodative—suggesting that differences between the two composite variables, mode of problem solving and inclusiveness, may not be as pronounced. The positioning of many countries is very similar to that presented quantitatively by Tosun and Howlett (2022) and qualitatively by others. For example, Greece scores low on everyone's classification scheme who has examined it (Tosun & Howlett, 2022; Zahariadis et al., 2021). Scholarship agrees that its style is clearly administrative. Howlett and Tosun (2019), Oztig (2022), and Tosun and Howlett (2022), and this article agree that Turkey similarly has a policy style with low capacity and low inclusiveness. Sweden lies at the other end of every study's scheme. There is widespread agreement that it has a managerial policy style (Zahariadis et al., 2021).

Differences raise more interesting points. They do not necessarily point to diverging perspectives but rather suggest areas for further empirical analysis. The UK and Germany differ

in policy styles, according to Richardson (1982). In an update to the British style, the same author (2018) notes a change in Britain's style. Complementing Howlett and Tosun (2019), our article finds that the change has made the two countries more similar. Exadaktylos (2022) argues the UK has an adversarial style (high capacity but low inclusiveness), but our data suggest otherwise in inclusiveness, 6.83, which is reasonably high relative to many of its peers in Group A. Interestingly, Tosun and Howlett (2022) find Germany's style to be more accommodative (low capacity and high inclusiveness), but in our classification, it scores high enough to be in the managerial cell along with the UK. The Netherlands scores on the managerial cell in Tosun and Howlett (2022) but on the accommodative cell in Richardson (1982) and ours. Further empirical research is needed to disentangle and clarify these issues.

The biggest surprises and divergence from other classificatory schemes come from France and the USA because they end up in opposite cells. France appears in the same cluster as Greece in Tosun and Howlett (2022), as having an administrative style. However, our classification expectations place it in the managerial style cell. Although it has one of the lowest scores of inclusiveness in its group, 6.33, its capacity for problem-solving is high, 6.69. Our classification expectation also differs substantially from that of Richardson (1982). The USA case is equally interesting. Whereas others classify it as having an adversarial style—for example, Taylor et al. (2002) and Howlett and Tosun (2019), although the latter do not label it as such—we find it diverges significantly from that picture, scoring low on both capacity and inclusiveness, 5.69 and 5.00, respectively. It has an administrative style similar to that of Eastern and Mediterranean European countries. In contrast, Tosun and Howlett (2022) find that it has what Zahariadis et al. (2021) call a managerial style. While more empirical studies are needed to elucidate the difference in expectations, one explanation may be that the USA is one of the countries that have experienced a marked deterioration in SGI indicators since 2014 (Tosun & Howlett, 2022, p. 171). It is possible that differences in individual country styles depend on the year covered by the data.

We report the scores of our composite variables for the 41 countries in our sample in Appendix: Table A1.

6 | CONCLUSION

Despite broad interest in national policy styles, the use of the concept has been hampered by a lack of quantitative operationalization. Tosun and Howlett (2022) productively steer the literature toward using "good governance" indicators from SGI. While accepting the core of Tosun and Howlett's (2022) argument, we have charted a different quantitative operationalization course, amending and advancing their claim. Still using SGI, we proposed using composite variables which more accurately capture the multiple dimensions of the concept of national policy style. We urge subsequent research work to validate, extend or amend our argument.

While we have addressed important methodological issues, we also urge caution when moving forward. Perhaps because of their conceptual richness, our composite variables highly correlate with one another ($r = 0.827$). This may pose problems of collinearity when using both variables in the same regression equation. We are aware results may vary in different settings and for this reason, we urge caution and suggest that analysts who encounter high collinearity employ different techniques, assuming policy style is not a control variable. If there is high collinearity and because using both at the same time can be problematic, analysts may either use each composite variable separately and then compare equation results or employ more

advanced techniques such as LASSO or partial least squares regression to address multicollinearity issues.

We acknowledge the composite variables that we constructed have strengths but also limitations. The main strength is their ability to capture a multi-dimensional concept, policy style, and express it in an easily interpretable summary form. However, apart from limitations inherent in the SGI data set (Croissant & Pelke, 2022), our argument is limited in two ways. First, the composite variables are proxies; SGI data were not collected for this purpose, and they may therefore measure different concepts in different combinations. Second, by being summary measures, composite variables lose some of the crispness and effect of their component indicators. For example, finding that policy style has an effect on, say, defense budget allocations, does not tell us whether this is due to policy-making or implementation. Studies offering policy advice may wish to revisit the individual components and compute individual effects to disentangle reasons for overall country policy performance (Freudenberg, 2003, p. 27).

Despite limitations, our argument points to a productive way forward. We helped operationalize the concept of policy style, thereby giving it greater analytical weight. We urge scholars to take the next step and test it quantitatively across many countries over a reasonably long period of time. Quantitative assessment of the effects of policy styles will surely yield additional insight into long-term national policy-making processes and outcomes.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

ORCID

Nikolaos Zahariadis http://orcid.org/0000-0003-2457-0561
Jörgen Sparf https://orcid.org/0000-0002-0202-0609
Evangelia Petridou https://orcid.org/0000-0001-7316-4899

REFERENCES

Balla, S. J. (2015). Political control, bureaucratic discretion, and public commenting on agency regulations. *Public Administration*, *93*(2), 524–538. https://doi.org/10.1111/padm.12137

Beavers, A. S., Lounsbury, J. W., Richards, J. K., Huck, S. W., Skolits, G. J., & Esquivel, S. L. (2013). Practical considerations for using exploratory factor analysis in educational research. *Practical Assessment, Research, and Evaluation*, *18*(1):6. https://doi.org/10.7275/qv2q-rk76

Bertelsmann Stiftung. (2022). *About the SGI*. Retrieved October 1, 2022, from https://www.sgi-network.org/2022/Mission_Statement

Cain, M. K., Zhang, Z., & Yuan, K. H. (2017). Univariate and multivariate skewness and kurtosis for measuring nonnormality: Prevalence, influence and estimation. *Behavior Research Methods*, *49*, 1716–1735. https://doi.org/10.3758/s13428-016-0814-1

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, *56*(2), 81–105. https://doi.org/10.1037/h0046016

Conway, J. M., & Huffcutt, A. I. (2003). A review and evaluation of exploratory factor analysis practices in organizational research. *Organizational Research Methods*, *6*(2), 147–168. https://doi.org/10.1177/F1094428103251541

Croissant, A., & Pelke, L. (2022). Measuring policy performance, democracy, and governance capacities: A conceptual and methodological assessment of the sustainable governance indicators (SGI). *European Policy Analysis*, *8*, 136–159. https://doi.org/10.1002/epa2.1141

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*(4), 281–302. https://doi.org/10.1037/h0040957

Everitt, B. S., Landau, S., Leese, M., & Stahl, D. (2011). *Cluster analysis*. Wiley.

Exadaktylos, T. (2022). Of "herd immunity" and inoculation investment: The british response to covid-19. In N. Zahariadis, E. Petridou, T. Exadaktylos, & J. Sparf (Eds.), *Policy styles and trust in the age of pandemics global threat, national responses*. Routledge.

Fabrigar, L. R., & Wegener, D. T. (2012). *Exploratory factor analysis*. Oxford University Press.

Ferreira, L., & Hitchcock, D. B. (2009). A comparison of hierarchical methods for clustering functional data. *Communications in Statistics—Simulation and Computation*, *38*(9), 1925–1949. https://doi.org/10.1080/03610910903168603

Freudenberg, M. (2003). Composite indicators of country performance: A critical assessment. *OECD Science, Technology and Industry Working Papers*, No. 2003/16, OECD Publishing, Paris: OECD. https://doi.org/10.1787/405566708255

Gordon, A. D. (1999). *Classification*, Chapman and Hall/CRC.

Gorsuch, R. L. (2015). *Factor analysis*. Routledge.

Howlett, M., & Tosun, J. (Eds.). (2019). *Policy styles and policy-making: Exploring the linkages*. Routledge.

Leiren, M. D., Inderberg, T. H. J., & Rayner, T. (2021). Policy styles, opportunity structures and proportionality: Comparing renewable electricity policies in the UK. *International Political Science Review*, *42*(1), 33–47. https://doi.org/10.1177/0192512120907112

Mazey, S., & Richardson, J. J. (1993). Conclusion: A European policy style? In S. Mazey & J. J. Richardson (Eds.), *Lobbying in the European community* (pp. 246–259). Oxford University Press.

Nardo, M., Saisana, M., Saltelli, A., & Tarantola, S. (2005). *Tools for composite indicators building*. https://publications.jrc.ec.europa.eu/repository/handle/JRC31473.

Oztig, L. I. (2022). Policy styles and pandemic management: The case of Turkey. *European Policy Analysis*, *8*(3), 261–276. https://doi.org/10.1002/epa2.1155

Peters, B. G. (2015). Policy capacity in public administration. *Policy and Society*, *34*(3–4), 219–228.

Pierson, P. (2004). *Politics in time: History, institutions, and social analysis*. Princeton University Press.

Richardson, J. (Ed.). (1982). *Policy styles in western Europe*. Routledge.

Richardson, J., Gustaffson, G., & Jordan, G. (1982). The concept of policy style. In J. Richardson (Ed.), *Policy styles in western Europe* (pp. 1–16). Routledge.

Schiller, C., Hellmann, T., & Paulini, P. (2022). *SGI Codebook: Sustainable governance indicators*. Bertelsmann Stiftung. Retrieved October 2, 2022, from https://www.sgi-network.org/docs/2022/basics/SGI2022_Codebook.pdf

Schraad-Tischler, D., & Seelkopf, L. (2015). *Concept and methodology— sustainable governance indicators 2015*. Bertelsmann Stiftung. Retrieved October 2, 2022, from https://www.sgi-network.org/docs/2020/basics/SGI_Concept.pdf

Sustainable Governance Indicators (SGI). (2022). *Evidence-based instruments*. Retrieved October 2, 2022, from https://www.sgi-network.org/2020/Good_Governance/Executive_Capacity/Evidence-based_Instruments/Quality_of_RIA_Process

Taylor, C., DeLeo, R. A., Crow, D. A., & Birkland, T. A. (2002). The US response to the covid-19 pandemic: Incoherent leadership, fractured federalism, and squandered capacity. In N. Zahariadis, E. Petridou, T. Exadaktylos, & J. Sparf (Eds.), *Policy styles and trust in the age of pandemics global threat, national responses*. Routledge.

Tosun, J., & Howlett, M. (2022). Analyzing national policy styles empirically using the Sustainable Governance Indicators (SGI): Insights into long-term patterns of policy-making. *European Policy Analysis*, *8*(2), 160–177. https://doi.org/10.1002/epa2.1142

Watkins, M. W. (2018). Exploratory factor analysis: A guide to best practice. *Journal of Black Psychology*, *44*(3), 219–246. https://doi.org/10.1177/0095798418771807

Wu, X., Ramesh, M., & Howlett, M. (2015). Policy capacity: A conceptual framework for understanding policy competences and capabilities. *Policy and Society*, *34*, 165–171. https://doi.org/10.1016/j.polsoc.2015.09.001

Zahariadis, N., Petridou, E., Exadaktylos, T., & Sparf, J. (2021). Policy styles and political trust in Europe's national responses to the COVID-19 crisis. *Policy Studies*, *44*(1), 46–67. https://doi.org/10.1080/01442872.2021.2019211

Zahariadis, N., Petridou, E., Exadaktylos, T., & Sparf, J. (Eds.). (2022). *Policy styles and trust in the age of pandemics global threat national responses*. Routledge.

How to cite this article: Zahariadis, N., Karokis-Mavrikos, V., Exadaktylos, T., Kyriakidis, A., Sparf, J., & Petridou, E. (2023). Advancing the operationalization of national policy styles. *European Policy Analysis*, 1–19. https://doi.org/10.1002/epa2.1172

APPENDIX

TABLE A1 Composite variables and policy style classifications.

Policy style	Country	Mode of problem solving	Inclusiveness
Administrative	Austria	5.72	6.17
	Bulgaria	4.78	4.83
	Croatia	4.11	4.17
	Cyprus	3.61	4.33
	Czechia	5.31	5.83
	Greece	5.56	4.00
	Hungary	4.81	3.83
	Iceland	5.78	6.00
	Italy	4.72	4.33
	Malta	5.78	6.17
	Poland	5.50	4.33
	Portugal	6.22	6.00
	Romania	3.97	3.83
	Slovakia	4.42	4.83
	Slovenia	4.92	5.83
	Turkey	5.08	3.67
	USA	5.69	5.00
Adversarial	Belgium	6.94	5.17
	Chile	7.11	5.33
	Lithuania	6.97	6.00
Accommodative	Mexico	6.00	6.33
	Netherlands	6.22	7.33
Managerial	Australia	7.31	7.17
	Canada	7.97	8.00
	Denmark	8.33	8.50
	Estonia	6.94	6.67

(Continues)

TABLE A1 (Continued)

Policy style	Country	Mode of problem solving	Inclusiveness
	Finland	8.50	8.50
	France	6.69	6.33
	Germany	6.44	6.50
	Ireland	6.72	6.83
	Israel	6.36	7.17
	Japan	6.31	6.50
	Latvia	7.42	7.17
	Luxembourg	6.72	7.33
	New Zealand	8.14	7.50
	Norway	7.36	9.17
	South Korea	6.81	6.83
	Spain	7.14	6.50
	Sweden	7.53	8.67
	Switzerland	7.17	8.83
	UK	7.31	6.83

AUTHOR BIOGRAPHIES

Nikolaos Zahariadis is the Mertie Bukman Chair and Professor of International Studies at Rhodes College, Memphis, TN, USA.

Vassilis Karokis-Mavrikos is a PhD Candidate in European Public Policy at the University of Surrey. His PhD thesis studies agenda-setting and policy formulation dynamics at the European Union level through a new EU-MSF lens. His research interests include Public Policy Change Drivers, Policy Entrepreneurship, Policy Styles, Public Health, and Digital Governance. He has published in prominent peer-reviewed journals and contributed to various edited volumes.

Theofanis Exadaktylos is Professor of European Politics at the University of Surrey, UK. His research interests focus on European public and foreign policies, Europeanization, the politics of austerity and crises, political trust and policy implementation, policy styles and Greek politics. He recently co-edited the book "Policy styles and trust in the age of pandemics: Global threat, national responses" (Routledge) with Nikolaos Zahariadis, Evangelia Petridou and Jörgen Spart. He is the co-editor of the Journal of Common Market Studies Annual Review and the convenor of the ECPR Standing Group on Political Methodology.

Alexandros Kyriakidis is Postdoctoral Research Fellow at Luiss Guido Carli University in Rome, Italy and Head of Operations & Research at the Center for Research on Democracy & Law, University of Macedonia, Thessaloniki, Greece.

Jörgen Sparf is Associate Professor in sociology, a founding member of the Risk and Crisis Research Center at Mid Sweden University and director of the immersive simulation environment RCR Lab. His theoretical interest is in relational aspects of institutions and organizations which he applies in the empirical field of societal resilience and disaster risk reduction. Some relevant journals that Jörgen has published in are *Policy and Society*, *Safety Science*, and *Journal* of *Contingencies* and *Crisis Management*.

Evangelia Petridou is associate professor of political science at Mid Sweden University, Sweden, and senior researcher at NTNU Social Research in Norway. She is co-editor of the International Review of Public Policy (IRPP). Her research interests center on policy process theories and specifically policy entrepreneurship, collaboration mechanisms in bureaucracies, and crisis management as a policy sector.