Disaggregating state fragility: a method to establish a multidimensional empirical typology

Jörn Grävingholt, Sebastian Ziaja & Merle Kreibaum

To cite this article: Jörn Grävingholt, Sebastian Ziaja & Merle Kreibaum (2015) Disaggregating state fragility: a method to establish a multidimensional empirical typology, Third World Quarterly, 36:7, 1281-1298, DOI: 10.1080/01436597.2015.1038340

To link to this article: http://dx.doi.org/10.1080/01436597.2015.1038340

© 2015 UNU-WIDER. Published by Taylor & Francis

Published online: 21 Jul 2015.

Article views: 444

View related articles

View Crossmark data
Disaggregating state fragility: a method to establish a multidimensional empirical typology

Jörn Grävingholt\textsuperscript{a*}, Sebastian Ziaja\textsuperscript{b} and Merle Kreibaum\textsuperscript{c}

\textsuperscript{a}Governance, Statehood and Security Department, German Development Institute (Deutsches Institut für Entwicklungspolitik, DIE), Bonn, Germany; \textsuperscript{b}Center for Distributional Conflict and Globalisation, Heidelberg University, Germany; \textsuperscript{c}Department of Economics, Göttingen University, Germany

(Received 31 March 2015; final version received 2 April 2015)

This conceptual and methodological article makes the case for a multidimensional empirical typology of state fragility. It presents a framework that defines fragile statehood as deficiencies in one or more of the core functions of the state: authority, capacity and legitimacy. Unlike available indices of state fragility, it suggests a route towards operationalisation that maintains this multidimensionality. The methodology presented should help in future research to identify clusters of countries that exhibit similar constellations of statehood, whereby ‘constellation’ refers to the specific mix of characteristics across the three dimensions. Such an identification of empirical types would fulfil a demand that exists both in academic research and among policy circles for finding a more realistic model of fragility at an intermediate level between single-case analyses and the far-too-broad category of state fragility.

Keywords: fragile states; statehood; fragility; state building; typology; mixture models

Introduction

The weakness, fragility or failure of states has evolved into one of the major narratives of politics and international relations in the post-cold war era. State fragility is assumed to have a profound impact on how key issues of global concern, such as climate change, poverty and violent conflict, can be addressed.\textsuperscript{1}

Since as early as the mid-1990s state failure has been a topic of concern in the pages of \textit{Third World Quarterly}. In a 1996 article Jean-Germain Gros developed a taxonomy of failed states, distinguishing anarchic, phantom, anaemic, captured and aborted states.\textsuperscript{2} This distinction had no lasting effect on the debate, however, and was never fully operationalised. Twelve years on, Charles Call attempted to
end the debate about state failure with the conclusion that the terminology was analytically useless and even harmful, because it ‘fuelled a tendency towards single, technocratic formulas for strengthening states, which emphasise coercive capabilities’. Call proposed abandoning the concept of state failure and all related terminologies, not least because of their state-centric implications.3

Yet the policy world began to breathe new life into the debate, diplomatically adopting the slightly less offensive term, ‘fragility’. Major development actors launched initiatives to analyse the character, consequences and policy implications of state fragility from a developmental perspective. As a consequence, state fragility – often coupled with violent conflict – has received high visibility in recent development policy documents, such as the European Report on Development 2009, a 2011 policy guidance paper by the OECD Development Assistance Committee, and the World Development Report 2011.4 At the Fourth High-Level Forum on Aid Effectiveness in Busan in November 2011 41 countries and multilateral organisations, including 17 fragile states, adopted a ‘New Deal on Engagement in Fragile States’. The prominent place the document was granted at this international forum underscored the perception of state fragility as a major barrier to achieving development objectives.5 Above all, however, these reports demonstrate a significant demand on the part of policy makers in international development for orientation in dealing with fragile states.

A rising number of fragility indices that have emerged over the past few years have tried to provide some of this orientation. Examples include the Fragile States Index (formerly ‘Failed States Index’), the Index of State Weakness, the State Fragility Index, the Political Instability Index and many others.6 All of them, however, share one weakness: they simplify the complicated reality behind the stability or decay of statehood to such an extent that they are of very limited use for the operational task of crafting policies to counter state fragility. The main issue with these indices is not so much the ever-difficult challenge of measurement but rather their common conceptual assumption that such a multidimensional concept as statehood can be aggregated and projected onto a one-dimensional scale, thereby allowing different dimensions to compensate for each other, without a substantial distortion of information.7 As an illustration, such diverse countries as Haiti, Pakistan and Zimbabwe end up in close proximity to each other in the 2014 Fragile States Index, although the respective challenges they face are rather different in nature.8

In general the authors of fragility indices are well aware of the limitations innate to their instruments: some of them recommend their index as a crude ‘early-warning’ tool that warrants further analysis of any given case, and they caution against too far-reaching interpretations based on their data. Others emphasise explicitly that ‘some of the weakest states perform poorly across the board, whereas others exhibit weakness in just one or two functions’.9 No attempt, however, has yet been made to incorporate the insight that state fragility is inherently multidimensional directly into the mapping exercise itself. Instead, authors only refer to the necessity to supplement their indices with detailed case-by-case analyses, thus de facto giving up on the opportunity to find larger patterns.

Yet policy makers ask for more. They know that individual case studies alone, without the methodological advantage of full comparative analysis, are
unable to provide the larger pattern – which is why indices are so popular, despite their obvious limitations. Aid bureaucracies in particular usually develop and prepare their instruments, ie the tools they either finance or implement themselves, not on a case-by-case basis. Instead they develop sets of instruments, or tool boxes, to which they resort when programming individual country assistance strategies. As a result of budget constraints and other bureaucratic incentives, to which aid agencies are subject like most large organisations, it would be unrealistic to expect aid organisations to develop their tool boxes for fragile states in every case as if ‘from scratch’.

From an academic point of view strict case-by-case approaches are not desirable either. Every external policy intervention in fragile states is based on assumptions about causes and effects, representing lessons that actors at different levels of the policy-making system draw from other cases which are deemed comparable. Obviously the quality of such learning depends significantly on the degree to which the cases used as points of reference are indeed sufficiently similar to the problem at hand. The very concept of ‘fragile states’ derived its traction from the compelling argument that aid to this group of countries had to be planned and delivered differently from development assistance to other places in the world. While this message was important, sufficient evidence exists today that fragility itself is far too broad a term, and encompasses too broad a range of countries and contexts, to allow for useful generalisations across the full population of this category – almost irrespective of how exactly the term is defined. However, if strict case-by-case decisions are logically not a way out of this conundrum, the inevitable challenge for policy-oriented research is to establish better categories of countries or contexts that allow for more suitable comparisons. This paper contributes to this task by suggesting a methodology for a data-driven typology of state fragility. As the underlying concept of fragility is multidimensional, and the typology would be derived from empirical data (rather than ‘ideal-typical’ theorising), we refer to this approach as a ‘multidimensional empirical typology’ of state fragility. The methodology presented should help in future research to identify clusters of countries exhibiting similar constellations of statehood, where ‘constellation’ refers to the specific mix of characteristics across a set of core dimensions of statehood.

Such a typology can then form the basis for further research into causal mechanisms of state weakness and external development interventions at a level of analysis that unpacks the broad category of fragility without sacrificing the heuristic advantage of comparison across cases. By introducing an ‘intermediate’ level of generalisation between the single case and the broad category of fragility, we do not mean to deny the importance of case-specific analyses. Rather, we assume that farsighted policies require analytical instruments that order reality according to both well founded theoretical assumptions and empirical observations. A more precise grasp of existing, typical forms of fragility should help development agencies and other actors to better prepare for the types of real-world situations they are most likely to be confronted with.

The article is structured as follows. The first section reviews the conceptual debate on state fragility, concluding that the multidimensional character of statehood should be taken more seriously, and specifies methodological conditions of data handling required to derive an empirical typology. In a second section we
argue that three functions in particular can be considered core dimensions of statehood and introduce attributes characterising each of them. The third section presents innovative methods of how empirical data could be used to first operationalise the measurement of each dimension of statehood and then identify clusters, or ‘real’ (i.e. empirical) types of state fragility. In a concluding section we discuss potential applications for such an empirical typology.

The conceptual debate on state fragility

The multidimensional character of statehood

All further debate depends, obviously, on the conceptual definition of state fragility. A closer look at most of the literature cited above reveals that the actual concept in question is not so much the fragility of the state as such in a legal or even ontological sense as it is the state’s ability to fulfil its basic functions, something that is also denoted as statehood. (This is why, in this paper, we use the concepts of state fragility and fragile statehood interchangeably.) As the state’s basic functions depend on its interaction with a given society, statehood is all about state–society relations.

Fragile statehood is characterised by a wide range of dysfunctional state–society relations – states not being able to control their territory, states not being able to support their population, and states failing to convince the population that they have legitimate claims to dominate a given territory. Thus, fragility is to some degree the antipode of state strength, but the focus is on state–society relations, and not on a state’s ability to fight other states. If fragility denotes the negative occurrence of statehood, the positive one is usually referred to as resilience. Empirically, however, fragility and resilience are not binary antipodes but only two ends of a logical field.

Both in its academic and its policy-oriented strands the literature on state fragility abounds in studies investigating the causes and consequences of fragility and internal conflict and discussing possible contributions towards making states more resilient. Parts of this literature are of a more general, conceptual nature. Others rely on in-depth case studies. Yet others use cross-country data to compute correlations and infer causality. In part to assist these latter efforts, and in part in order to satisfy public interest in easily accessible overviews of the ‘state of state fragility’ in the world, the plethora of indices that classify countries according to their statehood has emerged.

Despite the methodological weaknesses discussed in the introduction, most of these indices are in agreement with an increasing body of qualitative research that fragile states take very different forms and fragility should thus be analysed as a multi-faceted problem. Many authors would also argue that fragility problems – although varying in their scope and nature – revolve around a limited number of ‘functions’, ‘gaps’ or, as we prefer to refer to them, ‘dimensions’. Authors disaggregating fragility in two, three, four or more dimensions include Milliken and Krause; the Commission on Weak States and US National Security; Schneckener; Ghani et al; Patrick; Cliffe and Manning; Carment et al; Call; Stewart and Brown; and several others. All available attempts to map the degree of fragility of states in the world, however, have to date failed to incorporate this assumption of multidimensionality into their methodology.
By collapsing all data into one single measure per country (and, usually, year), most indices render the prior exercise of distinguishing different dimensions ex post useless.

The case for keeping dimensions distinct

A few recent contributions have highlighted the importance of disaggregating state fragility into key dimensions that need to be considered in their own right as well as in their interaction. Carment et al, proposing a three-dimensional view of statehood distinguishing between authority, legitimacy and capacity, present their Country Indicators for Foreign Policy not only in an aggregated form but also for each of the dimensions separately. Unfortunately the reader is left with the crucial task of structuring the resulting list of country scores in a meaningful way.

Similarly to Carment et al, Charles Call proposed considering ‘gaps’ in the areas of capacity, security and legitimacy as key factors driving state fragility (thus far relying on the Commission on Weak States and US National Security), but keeping them as distinct as possible rather than aggregating them in one measure. ‘Overlaps’ of two or three gaps in any given country should lead to interesting insights into their respective interaction, and policy responses in those cases should tackle the challenge of ‘balancing the gaps’. Call’s assertion is more than convincing. His model of intersecting circles, for instance, is well able to illustrate the difference between the exemplary cases mentioned above, Haiti and Zimbabwe.

Yet, when Call applies his model to the empirical world, his categorisation of states is based on an implicit binary logic: a gap exists or does not exist, depending on the arbitrary choice of a threshold. The consequence is two-fold: first, the classification of individual countries may be misleading because different thresholds might yield a completely different result. Second, the distinction of eight possible combinations is purely ideal-typical. It says nothing about the distribution of real-world phenomena along the three dimensions. In particular, this instrument is unable to shed light on the distribution of statehood in the broad ‘midfield’ of fragility. Yet it is precisely this midfield that policy makers should take an increasing interest in if they want to help prevent countries from slipping into a decay of statehood.

We therefore need an approach to mapping state fragility that fulfils two conditions. First, for every country and period of time, the approach measures each dimension of statehood separately along a continuum of values between ‘perfect’ and ‘fully defunct’. Second, the approach is able to identify clusters of countries that display ‘similar’ constellations of statehood across all measured dimensions.

One important consequence of this approach is obvious: it will not necessarily provide an answer to the question of which states should usefully be considered fragile, because the identification of types that differ along multiple dimensions will usually not allow the ordering of the resulting groups according to a joint fragility measure, and likewise, the ordering of countries within a group. The focus of this paper instead is on a methodology to identify patterns of statehood across countries rather than on labelling individual countries as fragile or not.
Three core dimensions of fragility (and statehood)

Which are the core dimensions that an analysis of state fragility should measure separately and treat as distinct characteristics of statehood? As we have seen in the previous section, different authors present several possible answers. However, the underlying concepts of state functions are not so dissimilar. As Call has observed, much of the literature has focused on two concepts: effectiveness and legitimacy.19 Others, such as Carment et al and Call himself, have argued that the provision of security is a capability that is different in nature from the delivery of services such as water supply and sanitation or primary schooling, since it is intimately related to the state’s ability to protect its authority vis-à-vis competing actors. These authors have thus proposed the distinction of three main dimensions.20 Cliffe and Manning, in turn, add public finance functions as an additional, analytically distinct fourth dimension (rather than viewing them as a subset of effective state capacity).21 In a similar vein Rice and Patrick propose economic, political, security and social welfare functions as the main pillars of statehood.22

In sum, there exists broad agreement on the attributes of statehood (and its lack); it is only the organisation into dimensions that is disputed. As most authors who measure fragility, however, do so by aggregating all data across all dimensions, the exact delineation of these very dimensions does not make a substantial difference. For our purpose, this is not true. We show that theoretical considerations speak in favour of a three-dimensional conceptualisation of statehood. Additionally, while success or failure on each dimension can have direct or indirect effects on the other dimensions, we argue that empirical cases show that none of these effects is automatic or linear.

In line with many of the authors discussed above, we propose to conceptualise statehood as a phenomenon that is constituted of three distinct, though interrelated, dimensions: state authority, capacity and legitimacy (ACL). This terminology is closest to those used by Carment et al, Call, and the Commission on Weak States.23 The exact focus of each of these dimensions is then derived from general conceptual considerations. Each of the three categories of authority, capacity and legitimacy represents a fundamental type of state–society relation and can be traced back to separate strands of political theory.

As Figure 1 illustrates, authority refers to the control of violence by the state, which, for this purpose, takes the freedom to use violence away from the
people. This coincides with a corporatist strand of political theory that views the state as a violence entrepreneur. It can be traced back to Thomas Hobbes’ idea of the state as the Leviathan, and was developed, among others, by Charles Tilly. More recently, North et al have contributed to this type of thinking. The authority dimension refers to the extent to which the state holds the monopoly of violence and can secure its claim on this monopoly against competitors. A diminished authority reduces the state’s ability to define and execute rules and protect citizens from wilful violence. By implication, authority is thus related to the degree that the state can guarantee the physical integrity of its citizens. However, a certain degree of violent crime seems to be unavoidable in any society, without necessarily calling the state’s monopoly of violence into question.

Examples of countries that exhibit particular challenges along the authority dimension include cases of civil war, violent insurgency or widespread criminal violence, such as Afghanistan and Iraq, but also some Central American countries.

State capacity, in turn, represents a state–society relationship that is characterised by the state giving basic services to the people. The exact scope of ‘basic’ services is not a definitional given, but can be assumed to include basic education, health care and a basic administration regulating social and economic activities sufficiently to increase collective gains and avoid massive negative externalities. Failure to perform in one or more of these areas diminishes the life chances of large parts of the population. The perspective of the state as a provider of services is that of a contractual relationship between state and society, as was developed by the 17th-century philosopher, John Locke.

Low state capacity is typical of many ‘classical’ developing countries, such as Senegal, Tanzania and Malawi, but also Haiti. Finally, legitimacy is about a type of state–society relationship in which society itself is active, in that it accepts, or refuses to accept, the state’s claim to be the only legitimate actor to set and enforce generally binding rules. As legitimacy is closely linked to the forging of a sense of identity within a society, this concept can be viewed as stemming from a constructivist perspective on the state. Legitimacy is notoriously complicated to measure. In line with our conceptualisation of legitimacy as acceptance of rule, we are interested in empirical rather than normative legitimacy. In theory, it should suffice to use mass surveys asking whether an individual accepts the rule exerted by the state authorities of their country as legitimate. In practice, the results of such surveys – if feasible in the first place – are that much less reliable the more illegitimate a state is. Any result would therefore suffer either from a systematic measurement error produced by exactly the property to be measured in the first place or at least from a strong theoretical assumption that such a bias should exist.

Countries with low levels of legitimacy certainly include some of the most repressive regimes in the world, such as North Korea, Uzbekistan and Turkmenistan, but low degrees of state legitimacy can also be the consequence of a breakdown of state services or of the provision of basic security.

Table 1 summarises our three-dimensional concept of statehood and the attributes related to each dimension. On the basis of this conceptualisation we can now turn to outlining a way of operationalising our proposed method.
Potential operationalisation

Our goal, as exposed above, is the identification of constellations of deficiencies in the core functions of the state, i.e., authority, capacity, and legitimacy. We conceptualised these core functions as latent traits, whose distribution will help identify existing constellations of fragility. We thus employ a two-stage procedure: the first stage generates continuous scores for each of the three dimensions; the second stage identifies fragility constellations by running a cluster analysis of this data.

Measuring the latent dimensions of state fragility

In order to obtain both reliable and valid scores for our latent dimensions, we suggest employing a small number of complementary indicators per dimension, transforming these indicators to reflect our theoretical expectations about the impact of the indicators on the dimension scores and aggregating the transformed indicators with a procedure that minimises measurement error.

The reason to settle for a small number of selected indicators is determined by the desire to reflect all the theoretical attributes of our fragility dimensions and by severe data limitations. As will be explained in detail below, it is difficult to identify more than a few indicators per dimension that validly represent our concepts of interest. We refrain from compromising validity by adding less satisfactory variables for the sake of reliability.²⁹

Opting for an approach that relies on a few indicators only, however, we have to choose an aggregation method that explicitly takes measurement error into account. Approaches that employ many indicators often argue that measurement error will cancel out on average.³⁰ Since we intend to use few indicators, which at the same time represent necessary conditions for each of their dimensions, we opt for a ‘minimum approach’: the ‘weakest link’, i.e., the smallest value among the indicators within each dimension, represents the best proxy for a country’s performance in a given year and is thus taken as the dimension score.³¹ As a result, our dimensions are immune to upward measurement error, unless it occurs in all indicators of one dimension at once.
The minimum approach also serves as a buffer against problems of missing data, as explained below.

The minimum approach requires that individual indicators are carefully aligned with each other, setting minimum and maximum values and the shape of the function between these extremes. A first step is the harmonisation of the direction of the indicators: lower scores correspond to deficiencies, higher scores to good performance. The second step is the standardisation of the indicators to a common scale – we propose one that ranges from 0 to 1. Setting the naturally occurring minima and maxima of potential indicators at 0 and 1 would not be practical, as different indicators might be considered to vary in their severity. Setting the endpoints with theoretical justification will also remove extreme outliers that would otherwise skew the distribution of the indicators. Being skewed may, however, remain a problem after setting the endpoints. Logging the variables is the most popular and usually sufficient transformation in order to mitigate this issue. In theoretical terms this corresponds to modelling decreasing marginal effects. Any transformation must retain or improve the conceptual meaning of the indicator, not decrease it.

Finally, whatever indicators are chosen, they must most probably be treated for missing data. Some types of lack can be treated theoretically. For example, indicators reporting comparatively rare events, such as battle deaths, can usually be assumed to be zero where no information is given. The missing data that remain must be addressed with statistical methods. Case-wise deletion is not an option, as it would significantly reduce the final sample and probably bias the results.

Instead, there are two options for dealing with incomplete information that one could apply here. The first is linear imputation and related approaches, such as moving averages. These techniques rely uniquely on the extant values in a country’s time series. They would be valid for the variables that can be expected to move rather slowly and with a high degree of path dependency. Nevertheless, this does not seem to be a valid approach if an indicator can be expected to react quickly to crises in a country. Nonetheless, the minimum approach would ignore the upward measurement bias of the wrongly imputed variable if any of the other indicators in a given dimension picked up the crisis and thus determined the dimension score.

A more sophisticated approach could possibly provide better replacements for missing data. So-called ‘multiple imputation’ incorporates a large number of additional socioeconomic variables (e.g., GDP per capita or inequality) for identifying correlations between them and our indicators of interest and thus for estimating missing information. While we do not implement the proposed typology here, the following sections discuss several options for indicators from existing datasets that could serve to measure the dimensions of state fragility, as well as making some suggestions for valid transformations. Issues such as the final selection of indicators, fine-tuning the transformations and calibrating the indicators among each other go beyond the scope of this paper.
**Authority**

As argued above, authority comprises the absence of competing claims to the monopoly of violence and the absence of large-scale occurrence of violence.

A promising option for measuring the former is the ‘monopoly of violence’ indicator from the Bertelsmann Transformation Index (BTI). It is based on expert judgement and ranges from 0 to 10. BTI data is available only for about 125 countries and only for every other year since 2006.

An expert assessment that offers wider coverage is the ‘political stability and the absence of violence’ indicator from the Worldwide Governance Indicators (WGI). This indicator is a meta-index of various expert assessments on trends in political violence. It is thus potentially more reliable than the BTI variant, which is based on the assessments of a few experts only, but not as good a match in terms of concept validity. The WGI’s methodology also sets the indicators’ averages to zero each year, thus removing global time trends.

Since experts may err, it is recommended to also draw on observational statistics. One option is battle-related deaths, taken from the database of the Uppsala Conflict Data Program. These include all casualties directly related to combat – civilians and military – on the territory of a specific country. This measure reflects the intensity of internal and external attacks on the integrity of a state and thus the degree to which the state faces organised (but only acute) challenges to its monopoly of violence. Whereas war size is usually defined by absolute battle deaths, it might be more suitable to employ battle deaths per capita for our purpose.

Another useful observational indicator is intentional homicides, ie ‘unlawful death purposefully inflicted on a person by another person’. The United Nations Office on Drugs and Crime (UNODC) collects information from international and national institutions on a yearly basis, stemming from the health or criminal justice sector. While individual instances of homicide do in the vast majority of cases not stem from explicit challenges to the dominance of the state, widespread lethal crime can be considered an indicator of organised crime in conflict with governing authorities or of a police force unable to stop rampant non-organised crime – both indicators of deficient state authority.

The range of authority issues is, of course, much wider than war and homicide. Other forms of violence, such as riots and guerrilla attacks, could add information to the dimension. One of the most prominent sources for this type of information is the Cross-national Time-series Data Archive. However, an incredibly low riot count for China, reaching a maximum of four riots per year during the past two decades, nurtures scepticism about the reliability of this source.

**Capacity**

Capacity comprises the ability of a state to provide its citizens with basic life chances. Useful indicators of protection from disease are access to improved drinking water sources, which is known to have a massive positive impact, as well as a low rate of under-five mortality per 1000 births. Eizenstat et al emphasise the merits of immunisation rates as indicators of general health policy efforts. Vaccination rates may, however, be easily biased, since religious or personal beliefs can lead to low compliance, despite a high ability of the state.
In terms of basic education, we would consider primary enrolment an appropriate measure, i.e. the ‘ratio of children of official school age [...] who are enrolled in school to the population of the corresponding official school age’. This variable is also available from the WDI database.

As for basic administration, the BTI offers an expert-coded indicator that measures the existence of fundamental structures of a civilian administration, such as a basic system of courts and tax authorities. While it does not assess their quality, a minimum degree of professionalism and meritocracy has to exist for a civilian state apparatus to be maintained. Just as with other BTI indicators, its limited coverage is problematic. Another indicator for administrative capacity is ‘bureaucracy quality’ from the International Country Risk Guide. However, the indicator is not freely available, which would limit the replicability of our model, and its methodology is not fully transparent, making validation an issue. In addition, it only ranges from 0 to 4, so that it does not offer a great deal of variation and, consequently, does not allow sufficiently for the differentiation of countries’ performances.

**Legitimacy**

As outlined above, legitimacy is difficult to measure. Lacking the nearly direct observations that reliable survey data would yield, we propose to use indirect measures based on the theoretical assumption that legitimate rule requires a lower degree of state repression to achieve obedience and drives fewer citizens out of their country for political reasons. With regard to the first aspect, the argument is that, because of its high cost, repression is only the second-best option for a state to resort to. Consequently a state will keep it at its lowest possible level: the less legitimate the state, the more it will depend on such measures.

One example of an indicator that measures repression is the Political Terror Scale (PTS). It assesses the degree to which a state resorts to violence (via ‘physical integrity violations’) in order to preserve its power. The drawback of this indicator is its ordinal scale, with only five levels. An indicator that – at first sight – offers a slightly more nuanced scale is the Cingranelli–Richards (CIRI) Human Rights Dataset’s indicator of physical integrity, which has a similar thrust to the PTS and uses the same raw data. However, CIRI aggregates repression across four equally weighted dimensions, which can theoretically result in extreme but one-dimensional repression being judged less severe than mild repression in several dimensions.

Our argument that a state which restricts media freedom has issues with legitimacy is based on the same reasoning as for physical integrity. Restricting media incurs costs and will only be attempted when free media would undermine the state’s ability to claim the support of the wider population. Reporters without Borders provide reasonable data on press freedom; however, since their change in methodology in 2012, they no longer offer a complete time series. An alternative measure, despite allegations of an ideological bias, is Freedom House’s Freedom of the Press indicator, as it covers a wide range of countries and is comparable across time.

In addition to degrees of repression, a more legitimate state can be expected to drive fewer citizens into political exile. From the population’s point of view,
even if they have no possibility of expressing their opinion publicly, they (usu-
ally) still have the option of ‘voting with their feet’, that is, of emigrating and
seeking asylum for political reasons, or, in Albert Hirschman’s famous terminol-
ogy, ‘exit’. In order to better distinguish political motivations for emigration
from other causes, we suggest considering only the number of granted
(not requested) asylums by country of origin. Again, absolute values should
be transformed into asylum seekers per capita.

Legitimacy and related concepts have also been measured with composite
indices, some of them with sufficient temporal and geographical reach. How-
ever, they usually incorporate strong normative components and bind legitimacy
closely to the concept of Western democracy. This is the case, for example with
‘legitimacy of the state’ from the Fund for Peace’s Fragile States Index or with
the WGI’s ‘voice and accountability’.

**Identifying fragility clusters**

Once the authority, capacity and legitimacy scores are generated, we will obtain
a three-dimensional data space with continuous variables ranging from 0 to 1,
as described above. Conventional fragility indices would now aggregated these
dimensions with arithmetic averaging. However, this blurs distinctions between
different fragility constellations, particularly in the middle ranges; different
dimensions are allowed to compensate for each other. For example, a country
that starts a civil war, thus decreasing its authority score, and at the same time
increases primary enrolment substantially, thus increasing its capacity score,
could end up with zero change on the aggregate in an index employing aver-
ages. In order to avoid this scenario, and to be able to distinguish countries
with different ‘gaps’, in the words of Charles Call, we suggest identifying clus-
ters with similar fragility issues instead.

In order to do so, we propose using a mixture model. Mixture models have
been used extensively in biology and medicine to derive (unobserved) classifi-
cations from (observable) properties such as size and colour of animals or symp-
toms and duration of illnesses. Only recently has clustering via mixture
modelling and related techniques become more prominent in political science.

This model-based clustering approach has various advantages over less for-
mal clustering, such as k-means or nearest neighbour clustering. Mixture models
provide information about the model fit and help select the best number of
groups to consider. For each observation they provide the probabilities of
belonging to any group, thus generating a measure of uncertainty that helps
improve the interpretation of the results compared to a deterministic classifi-
cation. While all these technical properties help identify a valid typology of states,
Justin Grimmer and Gary King remind us that the analytical utility of a classi-
ification should be regarded more highly than an ultimately arbitrary measure of
statistical fit.

One final aggregation issue that should be addressed is the time dimension. A
country’s condition is clearly dependent upon its condition in the previous year.
Modelling this dependence explicitly in a clustering exercise requires algorithms
that are not readily available. Calculating separate models for each year in the
sample could be one alternative. However, this complicates the comparison of
groups over time. It would require the definition of rules on when to consider a group equivalent to or distinct from a group in the previous year. The most conservative solution would be fixing time and pooling country years. This approach would detect those types of fragility that have materialised over the whole period under examination – not an overly bold assumption given an expected coverage of our main variables of less than 10 years. Furthermore, fixed groups make the analysis of movements of countries between groups over time more fruitful. Pooling country years would also be in line with a general assumption that typical constellations represent relatively stable equilibria across space and time.

**Conclusion**

In this article we have made the case for a multidimensional empirical typology of state fragility. We presented a conceptual framework derived from seminal theoretical works that defines state fragility as deficiencies in one or more of the core functions of the state: authority, capacity and legitimacy. We suggested a route towards operationalisation that maintains the multidimensional properties of our concept. Some initial suggestions for choosing appropriate indicators were provided.

Our main contribution to the literature on state fragility and the measurement of social science concepts in general is the emphasis on preserving conceptual multidimensionality in the operationalisation of concepts, thus achieving increased content validity.\(^58\) While most fragility indices insist that readers should consider individual dimension scores when using an index, it is usually only the aggregate index that receives substantial attention both in public and in academia. Given the current incentives in political science publishing, however, the step for an index from merely being available to being employed in quantitative regression analysis is a small one – all concerns about the validity of the measure notwithstanding.\(^59\)

The message of multidimensionality is also relevant for policy advice. The current practice of measuring fragility in one-dimensional scores is inherently unable to bring across the message that state fragility is in fact a multidimensional issue. Instead, the eye-catching ranking tables drown nuanced messages, no matter how much the latter are advertised. In addition, representatives of countries that rank low regularly consider these publications an insult instead of an incentive to improve. By contrast, distinguishing constellations of fragility provides much more information than simply defining a list of fragile-state countries. Development practitioners have long recognised that the issue of fragility is more complicated. An empirical typology fulfils the demand for a more realistic model of fragility without adding the complexity of complete disaggregation that would make the concept unmanageable. Information about which countries exhibit similar combinations of dysfunctional statehood will enable practitioners to pre-sort country cases for further analysis and intervention. The typology will not be able to provide an automatic recipe for a given case, but it will provide helpful diagnostics and meaningful pairings for cross-country comparisons on an intermediate level of analysis. Change over time will provide an additional level of information: countries that move from one group to another. The typology can serve to unveil substantial transformations of states, thus
either confirming existing knowledge or pointing at previously understudied cases that merit additional consideration. In a nutshell, a multidimensional empirical typology can make the much-maligned concept of fragility more useful for research and policy making.

Acknowledgments
We gratefully acknowledge comments and advice provided at various stages by Aaron Clauset, Jörg Faust, Rachel Gisselquist, Gary Goertz, Kristian Gleditsch, Joe Hewitt, Seth Kaplan, Andrew Mack, Imme Scholz, Stella Seibert-Palascino, Daniel Stegmüller, Mario Stumm and two anonymous reviewers, as well as by members of the World Bank’s IDA Working Group on Fragile and Conflict-affected Countries at a meeting in September 2011.

Funding
Research towards this paper was in part funded by the Federal Ministry for Economic Cooperation and Development (BMZ) under a research grant on Transformation and Development in Fragile States. The paper’s approach and findings are the sole responsibility of the authors and should in no way be taken to represent the views of BMZ or of any of the organisations and individuals mentioned above.

Notes on contributors
Jörn Grävingholt is Senior Researcher in the Governance, Statehood and Security Department, German Development Institute, Bonn. His research interests include state building and peace building, democracy promotion, the political economy of governance reforms and the role foreign aid. He has published articles in Democratization, Contemporary Politics, the Journal of International Relations and Development and various edited volumes.

Sebastian Ziaja is post-doctoral researcher at the Research Centre for Distributional Conflict and Globalisation, Heidelberg University. He obtained his PhD in 2014 from the Department of Government at the University of Essex with a thesis entitled “A Nudge too Far? Democracy Aid, Democratization and Political Instability.” His work on state fragility has been published by the United Nations Development Programme, the Journal for International Relations and Development and the Zeitschrift für vergleichende Politikwissenschaft.

Merle Kreibaum is a research associate at the Department of Economics, University of Göttingen. Her research interests include state fragility, post-conflict development and policy evaluations.

Notes
1. The quality of causal links between state fragility and some of the global threats most commonly referred to, such as terrorism or international crime, is rather dubious. See Patrick, Weak Links.
2. Gros, “Towards a Taxonomy.”
5. These documents and the processes behind them have contributed substantially to a more nuanced understanding of the underlying conflicts in fragile states. The tendency of overly state-centric approaches that Call criticised is increasingly balanced by the recognition that the legitimacy of political institutions, justice and inclusion are key elements in overcoming state fragility.
6. For a comprehensive overview of state fragility indices, see Fabra Mata and Ziaja, Users’ Guide.
7. Ibid; Gutiérrez Sanín, “Evaluating State Performance”; Ziaja, “What do Fragility Indices Measure?”; and Gisselquist, “Developing and Evaluating Governance Indexes.” With respect to their widespread ‘use and abuse’, fragility indicators have followed in the footsteps of governance indicators, which were beset with similar problems. See Arndt and Oman, Users and Abuses.
10. For a discussion of the bureaucratic logic of aid organisations using the example of donor fragmentation, see Knack and Rahman, “Donor Fragmentation.” For an application of this idea to the context of fragile states, see Faust et al., “Foreign Aid,” 12–14.
12. Rotberg, State Failure; Paris, At War’s End; Schlichte, The Dynamics of States; and Call and Wyeth, Building States.
17. Call, “Beyond the ‘Failed State’.”
18. Giraudy takes a similar methodological approach to suggesting a typology of states by ‘state strength’. Giraudy, “Conceptualizing State Strength.”
19. Call, “Beyond the ‘Failed State’,” 305. It was also a broadly shared view among OECD donor agencies during much of the 2000s that fragile states were those that were either unable or unwilling to carry out key tasks.
21. Cliffe and Manning, “Practical Approaches.” In addition, Cliffe and Manning consider the provision of rule of law to be a crucial aspect closely linked to the security function (rather than to state capacity).
22. Rice and Patrick, Index of State Weakness.
26. OECD, Service Delivery.
30. Kaufmann et al., Governance Matters VIII.
31. Goertz and Mahoney, A Tale of Two Cultures, 31.
32. An example from the indicators described below could be that the highest homicide rate to be found anywhere in the world does not necessarily imply the same low degree of authority as the highest rate of battle deaths does.
33. Fabra Mata and Ziaja, Users’ Guide.
34. Goertz and Mahoney, A Tale of Two Cultures, 150–160.
37. Bertelsmann Stiftung, Bertelsmann Transformation Index, 73.
38. Kaufmann et al., Governance Matters VIII.
39. UCDP, Battle-related Deaths Dataset.
40. UNODC, Homicide Counts.
41. Banks and Wilson, Cross-national Time-series.
42. Eizenstat et al., “Rebuilding Weak States.”
44. Bertelsmann Stiftung, Bertelsmann Transformation Index, 74.
46. Wintrobe, The Political Economy; and Dogan, “Conceptions of Legitimacy,” 120.
47. Gibney et al., Political Terror Scale.
49. Wood and Gibney, “The Political Terror Scale.”
50. RSF, World Press Freedom Index.
51. Freedom House, Freedom of the Press. For a critical assessment, see Schneider, Media Freedom Indices.
52. Hirschman, Exit, Voice, and Loyalty.
Bibliography


