LAW AND MULTIDIMENSIONAL MEASUREMENT

Ronald J. Coleman & Ana Vaz

ABSTRACT

Without accurate measurement of phenomena, it can be difficult to enact targeted corrective policies. While measurement in the social sciences is always challenging, it becomes infinitely more so when the phenomena under study are multifaceted and, thus, cannot easily be assessed by reference to a single factor. Phenomena of this latter variety might include governance, empowerment, or rule of law. Multidimensional measurement techniques have sought to capture such phenomena. Legal scholarship utilizing multidimensional measurement has, to date, generally focused on techniques treating each factor independently, which fails to reveal the interdependency of different factors. This Article seeks to fill that methodological gap by presenting an alternative measurement framework, and illustrating its value for the legal community through a criminal justice index we have constructed.

I. INTRODUCTION

Without accurate measurement of phenomena, it can be difficult to enact targeted corrective policies. While measurement in the social sciences is always challenging, it becomes infinitely more so when the phenomena under study are multifaceted and, thus, cannot easily be assessed by reference to a single factor. Phenomena of this latter variety might include governance, empowerment, or rule of law. Multidimensional measurement techniques have sought to capture such phenomena. Legal scholarship utilizing multidimensional measurement has, to date, generally focused on dashboards[1] and composite indices.[2] While extremely useful, such techniques focus on

---

[1] *See, e.g.*, Daniel A. Farber, *California Climate Law—Model or Object Lesson?*, 32 PACE ENVTL. L. REV. 492 (2015) (climate regulation); *see also* Kevin N. Wright, *Developing a National Performance Measurement System*, 67 FED. PROBATION 37 (2003). The dashboard approach consists of simply measuring different facets of a phenomenon independently, without aggregating them.

[2] *See, e.g.*, Stephen J. Choi et al., *Judicial Evaluations and Information Forcing: Ranking State High Courts and Their Judges*, 58 DUKE L.J. 1313 (2009) (judicial evaluation); Charles H. Logan, *Well Kept: Comparing Quality of Confinement in Private and Public Prisons*, 83 J. CRIM. L. & CRIMINOLOGY 577 (1992) (prison quality); Jean-François Roberge, *"Sense of Access to Justice" as a Framework for Civil Procedure Justice Reform: An Empirical Assessment of Judicial Settlement*

each factor individually, and so fail to reveal how different factors are interdependent.[3] The purpose of this Article is to illustrate the value of creating multidimensional measures using a framework addressing this shortfall: the Alkire-Foster method (AF Method).[4]

The AF Method was initially developed to measure poverty across a range of factors beyond income, such as health, education, threats of violence, and disempowerment.[5] It has been used, among other things, in the development of a global multidimensional poverty index,[6] a women's empowerment in agriculture index,[7] Bhutan's gross national happiness index,[8] and national measures of poverty.[9] The AF Method is extremely flexible. It permits the selection of different indicators to create bespoke measures that fit particular contexts and are targeted to specific purposes.[10] Measures based on this method are easy to break down and appropriate for

Conferences in Quebec (Canada), 17 CARDOZO J. CONFLICT RESOL. 323 (2016) (access to justice). Like a dashboard, a composite index also begins by measuring the facets of a phenomenon independently. However, a composite index then aggregates these independent measurements into a single number. *See* SABINA ALKIRE ET AL., MULTIDIMENSIONAL POVERTY MEASUREMENT AND ANALYSIS 73-74 (2015) [hereinafter POVERTY MEASUREMENT].

[3]   *See* WORLD BANK GROUP, MONITORING GLOBAL POVERTY: REPORT OF THE COMMISSION ON GLOBAL POVERTY 168-171 (2017); POVERTY MEASUREMENT, *supra* note 2, at 74-75.

[4]   The Alkire-Foster method was created by Sabina Alkire, of the Oxford Poverty and Human Development Initiative (OPHI) at the University of Oxford, and Professor James Foster, of the Elliot School of International Affairs at George Washington University. *See generally* Sabina Alkire & James Foster, *Counting and Multidimensional Poverty Measurement*, 95 J. PUB. ECON. 476 (2011) [hereinafter Alkire & Foster]. It should be noted that, in addition to dashboards and composite indices, other multidimensional measurement approaches include Venn diagrams and statistical approaches (e.g., principal component analysis and multiple correspondence analysis). The AF Method is also preferable to these other techniques, because, for instance, unlike Venn diagrams it produces a summary measure and allows for the consideration of a relatively large number of factors, and unlike the statistical approaches, it permits easier comparison across metrics based on different data sets. POVERTY MEASUREMENT, *supra* note 2, at 70-122.

[5]   *Policy – A Multidimensional Approach*, OXFORD POV. & HUM. DEV. INITIATIVE, https://ophi.org.uk/policy/multidimensional-poverty-index/ (last visited July 19, 2019).

[6]   This index was developed by OPHI and the United Nations Development Program. *See* Sabina Alkire & Selim Jahan, *Assessing Multidimensional Poverty – One Index at the Global Level*, UNITED NATIONS HUM. DEV. PROGRAMME: HUM. DEV. REP. (Feb. 27, 2018), http://hdr.undp.org/en/content/assessing-multidimensional-poverty-%E2%80%93-one-index-global-level; *see also Multidimensional Measures in the Sustainable Development Goals*, OXFORD POV. & HUM. DEV. INITIATIVE, https://ophi.org.uk/ policy/multidimensional-measures-in-the-sustainable-development-goals-poverty-and-gross-national-happiness/ (last visited July 27, 2019).

[7]   This index was launched by the International Food Policy Research Institute with OPHI and the United States Agency for International Development. *See* Sabina Alkire et al., *The Women's Empowerment in Agriculture Index*, 52 WORLD DEV. 71 (2013).

[8]   *Bhutan's Gross National Happiness Index*, OXFORD POV. & HUM. DEV. INITIATIVE, https://ophi.org.uk/policy/national-policy/gross-national-happiness-index/ (last visited July 19, 2019).

[9]   *Who Uses a Multidimensional Approach?*, MULTIDIMENSIONAL POV. PEER NETWORK, https://www.mppn.org/multidimensional-poverty/who-uses/ (last visited July 19, 2019).

[10]   *How is the Alkire-Foster Method Used in Policy?*, OXFORD POV. & HUM. DEV. INITIATIVE, https://ophi.org.uk/policy/national-policy/ (last visited July 19, 2019).

various forms of data. Although the AF Method has been largely ignored in legal scholarship,[11] we believe the framework has myriad applications in the legal context,[12] and we will illustrate this by presenting a criminal justice measure we have constructed: the Law Enforcement Capacity Index (LECI). The remainder of this Article will proceed as follows: Section II will explore the AF Method, including how to create and utilize bespoke empirical measures; Section III will present our LECI and review results and analyses of such measure; and Section IV will conclude.

II. THE ALKIRE-FOSTER METHOD

The AF Method creates a multidimensional measurement framework.[13] This framework helps reveal the extent to which units under study fall below an established standard, i.e., the extent to which units (perhaps individuals or institutions) are inadequate according to some selected criteria. For instance, the framework may be used to estimate the quantum of individuals or households lacking access to justice, identify public agencies tainted by corruption, or perform impact evaluations of legal policies. Measures based on the AF Method may generate, among other things, statistics on the incidence,[14] intensity,[15] and composition[16] of the type of inadequacy under analysis.[17] The framework is particularly suited to supporting policy analysis because it: (i) produces an overall headline measure; (ii) allows for identification of the factors driving inadequacy; (iii) permits comparisons

11   *Cf.* Matthew D. Adler, *Equity by the Numbers: Measuring Poverty, Inequality, and Injustice*, 66 ALA. L. REV. 551, 572-73 (2015) (mentioning method strictly in poverty context); Michael P. Vandenbergh et al., *Micro-Offsets and Macro-Transformation: An Inconvenient View of Climate Change Justice*, 33 HARV. ENVTL. L. REV. 303, 336 n. 154 (2009) (citing paper on method in poverty-related context).

12   An added benefit of lawyers understanding this empirical technique is that such lawyers will be better prepared to cross-examine expert witnesses employing multidimensional analyses. *See, e.g.,* Franklin M. Fisher, *Multiple Regression in Legal Proceedings*, 80 COLUM. L. REV. 702, 736 (1980) (noting value to lawyers of understanding multiple regression in view of possible use of regression by an opponent).

13   This Section draws upon the AF Method framework, initially set out in Alkire & Foster, *supra* note 4, and further described in POVERTY MEASUREMENT, *supra* note 2.

14   *See* OXFORD POV. & HUM. DEV. INITIATIVE, CONSTRUCTING A MULTIDIMENSIONAL POVERTY MEASURE (2015), http://www.ophi.org.uk/wp-content/uploads/Constructing-a-multidimensional-poverty-index-WEB-Jan-2015.pdf. By "incidence of inadequacy," we mean the percentage of analyzed units that are inadequate. *Id.*

15   *See id.* By "intensity of inadequacy," we mean the average proportion of insufficiencies faced by inadequate units simultaneously. *Id.* The concept of insufficiency is discussed further *infra* Section II.A.

16   *See id.* By "composition of inadequacy," we mean the percentage of units that are both inadequate and deprived in each indicator.

17   For purposes of this Article, we will use the terminology "inadequate" and "insufficient," rather than "poor" and "deprived" (the terms normally used in the poverty literature), since our interest here is to reflect the measurement's capacity outside the poverty context. *See generally*, POVERTY MEASUREMENT, *supra* note 2.

across subgroups (e.g., by location, size, gender, or race); (iv) facilitates analysis of the evolution of inadequacy over time; and (v) is compatible with both cardinal and ordinal data.[18] This section will detail how to construct a measure using the framework and perform related analyses.

A. Constructing the Measure

Constructing a measure requires defining what one seeks to capture, setting the measure's parameters, and calculating the measure. The decisions on parameters often involve normative judgements, and so it is important to promote transparency by documenting the decision process and presenting the structure of the measure clearly.[19] Although there are not necessarily defined "steps" for creating a measure, for convenience, we present the process of creating a measure as consisting of several distinct steps.[20] These steps are summarized in Table 1.

The first step is to define the measure's purpose.[21] It is necessary to isolate the phenomenon one is seeking to study and the rationale for such study. For instance, a measure of quality of justice may be created to help guide budget or policy decisions, aid particularly underserved populations, monitor improvements over time, or complement other collected statistics.[22] It is critical to clearly establish the purpose for the measure at the outset, since this decision guides subsequent steps in constructing the measure.[23]

The second step is to establish the unit of identification.[24] The unit of identification is the entity identified by the measure as inadequate or adequate.[25] For instance, one might seek to analyze cities or states, certain

---

18	*Id.* at 21.

19	Some critics have attacked certain legal metrics as "oversimplifications of complex social phenomena that secretly reflect the unstated biases of their creators." Kevin E. Davis, *Legal Indicators: The Power of Quantitative Measures of Law*, 10 ANN. REV. L. & SOC. SCI. 37, 38, 48 (2014). Transparency in creating and presenting a measure may permit scrutiny of such biases.

20	We have set out these steps in the way we believe most beneficial for use by the legal academy and community. There are different and more granulated approaches than what we reflect here. *See, e.g.*, POVERTY MEASUREMENT, *supra* note 2, at 145-68.

21	*See id.* at 197-98 (discussing measure's purpose as an initial normative decision).

22	Of course, a measure by itself cannot generate favorable policy, but it may be designed with such a goal in mind. *See id.* at 20.

23	There are a great number of decisions to make in designing a measure. For instance, one may have to balance priorities, such as the measure being easy to describe, technically solid, operationally viable (e.g., having appropriate data), easily replicable, salient from a policy perspective, and fit to its stated purpose. *See id.* at 194-96. Having a clear sense of the measure's purpose at the outset is essential for navigating relevant tradeoffs. *Id.*

24	*Cf. id.* at 199-201.

25	*See id.* at 200.

courts, police precincts, or individuals. What should guide the choice of appropriate unit of identification is the measure's purpose.[26]

The third step is to select indicators[27] for inclusion in the measure and assign them weights.[28] When conceptualizing the phenomenon under analysis, a researcher might identify certain variables—such as income, educational attainment, or number of arrests—as most relevant to capture the phenomenon and select such variables as indicators. If helpful, a researcher may also group several conceptually similar indicators into a dimension,[29] such as by grouping three indicators relating to sufficiency of police staffing into a dimension on "staffing." In order to combine indicators into a measure, it is necessary to weight the indicators. Accordingly, each indicator is given a weight based on the importance of that indicator as compared to other indicators in the measure.[30]

The fourth step is to decide on insufficiency cutoffs for indicators and the overall inadequacy cutoff.[31] Here, we are concerned with what *minimums* must be met. The insufficiency cutoff for an indicator reflects the minimum attainment required so as not to be insufficient in such indicator.[32] The inadequacy cutoff reflects what minimum share of weighted insufficiencies would be necessary to identify a unit as inadequate.[33]

---

[26]   In some instances, it is relevant to draw a distinction between the unit of identification and the unit of analysis. *See id.* at 199-200, 220. This might take place when, for instance, data is available only at the household level, but one is interested in reporting results at the individual level. In that event, the unit of identification would be the household, and the unit of analysis would be the individual. *Id.* at 200, 220.

[27]   An indicator may be defined as "a data element that represents statistical data for a specified time, place, and other characteristics." *See Glossary of Statistical Terms: Statistical Indicator*, ORG. FOR ECON. CO-OPERATION & DEV., https://stats.oecd.org/glossary/detail.asp? ID=2547 (last visited July 19, 2019). Indicators may be considered the "building blocks of a measure [.]" *See* POVERTY MEASUREMENT, *supra* note 2, at 197.

[28]   *Cf.* POVERTY MEASUREMENT, *supra* note 2, at 210-11.

[29]   *Id.* Dimensions may be defined as "conceptual categories into which indicators may be arranged (and possibly weighted) for intuition and ease of communication." *See id.* at 197. Grouping related indicators into dimensions may aid communication of the measure's results, since there would normally be fewer dimensions than there are indicators, and the thematic dimensions may be more accessible to those less connected to the research. *Id.* at 202.

[30]   Put differently, the weight assigned to an indicator reflects the value that an insufficiency in such indicator has for inadequacy, relative to insufficiencies in other indicators. *See id.* at 197.

[31]   *Id.* at 197, 208-09.

[32]   *Id.* at 197.

[33]   *Id.* When building a measure, one has various options for inadequacy cutoffs with respective tradeoffs. For instance, one could deem a unit inadequate if such unit were insufficient in at least one indicator (called the union approach). Alkire & Foster, *supra* note 4, at 477. This approach would generally identify a large group of units as inadequate, *see* POVERTY MEASUREMENT, *supra* note 2, at 152, potentially including some which are only insufficient in a single indicator and whose performance may not be impaired by such insufficiency. An alternative option might be to deem a unit inadequate only if it were insufficient in all indicators (called the intersection approach). Alkire & Foster, *supra* note 4, at 477. This approach generally identifies as inadequate a very small group of units, perhaps leaving out units with many insufficiencies, *see* POVERTY MEASUREMENT, *supra* note 2, at 152, whose performance might be hindered even though they are not insufficient in all

The fifth step is to calculate the measure. This begins with identifying which units are inadequate. Suppose you have a population of n units and information on their attainments in d indicators. Let x_{ij} represent the attainment of unit i on indicator j. Assume w_j stands for the relative weight of indicator j, and the weights of the d indicators sum to one: $\sum_{j=1}^{d} w_j = 1$. Then, let z_j reflect the insufficiency cutoff for indicator j, and k denote the overall inadequacy cutoff. Unit i is identified as insufficient in indicator j if its attainment on that indicator is below the respective insufficiency cutoff: $g_{ij} = 1$ if $x_{ij} < z_j$ and $g_{ij} = 0$ if $x_{ij} \geq z_j$.[34] The inadequacy score of unit i, denoted c_i, is the weighted sum of its insufficiencies: $c_i = \sum_{j=1}^{d} w_j g_{ij}$.[35] Unit i is identified as inadequate if its inadequacy score is equal to or greater than the inadequacy cutoff: $c_i \geq k$.[36]

After identifying the inadequate units, one calculates the incidence of inadequacy (who is inadequate), intensity of inadequacy (how inadequate they are), and multidimensional index (a measure considering both the incidence and intensity).[37] The incidence of inadequacy (also referred to as the headcount ratio), denoted by H, is the proportion of inadequate units: $H = \frac{q}{n}$, where q is the number of inadequate units.[38] The intensity reflects the breadth of inadequacy and is the average inadequacy score among the inadequate units: $A = \frac{1}{q}\sum_{i=1}^{n} c_i I(c_i \geq k)$, where $I(.)$ is an identification function that assumes the value one if the condition between parentheses is true for unit i, and zero otherwise.[39] The multidimensional index (also referred to as the adjusted headcount ratio), denoted M_0, reflects the incidence of inadequacy adjusted for the intensity: $M_0 = HA$ or $M_0 = \frac{1}{n}\sum_{i=1}^{n} c_i I(c_i \geq k)$.[40]

The multidimensional index corresponds to the insufficiencies experienced by inadequate units expressed as a proportion of all possible insufficiencies (if all units were insufficient in all indicators).[41] In being sensitive to both the incidence *and* intensity of inadequacy, the multidimensional index can capture the effects of policies that either reduce the number of inadequate units or improve the position of inadequate units. For example, suppose a policy was successful at reducing the number of insufficiencies experienced by a set of highly inadequate units, but such

---

indicators. Where appropriate, it is helpful to select an inadequacy cutoff between these two extremes, *see id.*, potentially permitting one to identify as inadequate only those units with enough insufficiencies as might compromise a unit's performance.

[34]   Alkire & Foster, *supra* note 4, at 477-80.
[35]   *Id.*
[36]   *Id.*
[37]   *See* OXFORD POV. & HUM. DEV. INITIATIVE, *supra* note 14.
[38]   Alkire & Foster, *supra* note 4, at 477-80.
[39]   *Id.*
[40]   *Id.*
[41]   *Id.*

policy failed to make any inadequate unit adequate. A measure focused only on incidence would fail to reveal the value of such policy, but the multidimensional index would capture it.

Table 1: Multidimensional Measure Creation

Steps	Action	Short Description
Step 1	Define Measure's Purpose	Determine why the measure is being created
Step 2	Establish Unit of Identification	Determine the entity that will be identified as sufficient or insufficient
Step 3	Select Indicators and Assign Weights	Select the set of indicators that will be used, and assign such indicators individual weights
Step 4	Establish Cutoffs	Determine the insufficiency cutoffs for indicators and the inadequacy cutoff
Step 5	Calculate Measure	Identify and calculate the proportion of inadequate units (incidence or headcount ratio), breadth of inadequacy (intensity), and overall inadequacy (multidimensional index or adjusted headcount ratio)

B. Analyses of the Measure

After calculation of the multidimensional index, there are a number of additional analyses that may be made. We mention four such analyses below and summarize them in Table 2.

First, in order to study the pattern of insufficiencies among the population, one can estimate the uncensored and censored headcount ratios.[42] The uncensored headcount ratio of indicator j, denoted h_j, is the proportion of units that are insufficient in that indicator: $h_j = \frac{1}{n}\sum_{i=1}^{n} g_{ij}$.[43] The uncensored headcount ratios summarize the prevalence of the different insufficiencies among the population.[44] The censored headcount ratio of indicator j, denoted $h_j(k)$, is the proportion of units that are inadequate and insufficient in that indicator: $h_j(k) = \frac{1}{n}\sum_{i=1}^{n} g_{ij}I(c_i \geq k)$.[45] The censored headcount ratios have value, in that they summarize the prevalence of insufficiencies experienced by only the inadequate units.[46]

Second, it is possible to investigate the drivers of inadequacy. The measure can be broken down by the contribution of each indicator and

---

[42]   *See* POVERTY MEASUREMENT, *supra* note 2, at 165-67.
[43]   *Id.*
[44]   *Id.*
[45]   *Id.*
[46]   Please note that, by definition, for any given indicators, the censored headcount ratio is always smaller than, or equal to, the uncensored headcount ratio.

dimension.[47]   Because the multidimensional index can be written as the weighted sum of the censored headcount ratios ($M_0 = \sum_{j=1}^{d} w_j\, h_j(k)$), the relative contribution of an indicator is obtained by multiplying the indicator's censored headcount ratio by the indicator's weight and dividing by the index.[48]   A dimension with a high relative contribution might become a policy priority.[49]   For example, suppose a policymaker aims to reduce inadequacy in access to justice, and the created measure reflects that 40% of the inadequacy in the measure derives from an indicator for cost of legal services, while the other eight indicators account for 10% or less each. In such circumstances, a policymaker might most easily reduce inadequacy by taking actions targeted at the cost of legal services, such as improving dissemination of information regarding pro bono legal services.

Third, the measure may be decomposed by subgroups, such as location, size, age, or gender.[50]   Suppose the population can be divided into m exhaustive and mutually exclusive subgroups, $M_0{}^l$ is the multidimensional index for subgroup l, and v^l denotes the population share of such group. Then, the multidimensional index can be expressed as the weighted sum of the subgroups' multidimensional indices: $M_0 = \sum_{l=1}^{m} v^l M_0{}^l$.[51]   This feature allows one to analyze the situation of particular subgroups or draw comparisons between the performances of different subgroups. Subgroup results may inform policy,[52] for instance, by helping target resources to the groups most in need. Combining the subgroup decomposition with the breakdown by indicators permits display of the composition of inadequacy by each subgroup.[53]   Returning to the example of access to justice, suppose the measure reflected that an indicator for linguistic barriers, rather than cost of legal services, drives inadequacy among a specific ethnic minority. In such circumstances, a legal aid organization seeking to increase access to justice for that ethnic minority might be best served by, for instance, increasing use of interpretation and translation services.

Fourth, and finally, if there is comparable data for different time periods, it is possible to track changes in the measure over time to investigate whether inadequacy and its partial indices—incidence and intensity—are decreasing or increasing.[54]   It can be particularly interesting to analyze

---

[47]   POVERTY MEASUREMENT, *supra* note 2, at 166, 185.

[48]   *Id.* at 185.

[49]   *Id.* at 186-87; *see also* Jeffrey K. Staton, *A Comment on the Rule of Law Unplugged*, 59 EMORY L.J. 1495, 1513 (2010) (highlighting value of subdimension analysis for certain legal phenomena).

[50]   POVERTY MEASUREMENT, *supra* note 2, at 184.

[51]   *See* Alkire & Foster, *supra* note 4, at 480.

[52]   POVERTY MEASUREMENT, *supra* note 2, at 185-87.

[53]   *Id.* at 186-187.

[54]   To do so, for each of the outcomes of interest (e.g., inadequacy, incidence, intensity, and censored headcount ratios) one calculates the absolute and the relative rates of change. *See* Sabina Alkire, Jose Manuel Roche, & Ana Vaz, *Changes Over Time in Multidimensional Poverty: Methodology*

changes in inadequacy over time broken down by subgroups, as this can reveal whether certain subgroups are increasing or reducing inadequacy faster than the overall population.

Table 2: Selected Analyses

Analysis	Short Description
Uncensored and Censored Headcount Ratios	Calculate the proportion of units insufficient in each indicator (uncensored headcount ratios) and the proportion of units that are both inadequate and insufficient in each indicator (censored headcount ratios)
Percentage Contribution	Compute each indicator's contribution to inadequacy
Decomposition	Disaggregate the index by subgroups (e.g., location, size, age, gender) to analyze any relevant concentrations of inadequacy
Estimating Changes Over Time	When data from different periods of time is available, estimate the rates of change

III. EMPIRICAL APPLICATION:
LAW ENFORCEMENT CAPACITY INDEX

For illustrative purposes, we have created a measure of law enforcement capacity utilizing survey data from the 2013 Law Enforcement Management and Administrative Statistics (LEMAS) study, authored by the U.S. Department of Justice, Bureau of Justice Statistics and produced by the Inter-university Consortium for Political and Social Research.[55] We refer to our measure as the Law Enforcement Capacity Index or "LECI."[56] Importantly, while we believe the measure may have some substantive value, it is

---

and Results for 34 Countries, 94 WORLD DEV. 232, 232-34 (2017). The "absolute" rate of change is simply the difference in level of the outcome of interest between two different time periods. *Id.* The "relative" rate of change is the difference in levels of the outcome across two different periods defined as a percentage of the level observed in the first period. *Id.* For example, suppose that the incidence of inadequacy was 20% in year 2017 and 40% in year 2018. Based on this information, we could say that the rate of absolute change of incidence was 20 percentage points (i.e., 40-20), while the rate of relative change was 100% (i.e., (40-20)/20). *See id.*; *see also* POVERTY MEASUREMENT, *supra* note 2, at 265.

[55] *See* U.S. DEP'T OF JUSTICE, OFFICE OF JUSTICE PROGRAMS, BUREAU OF JUSTICE STATISTICS, LAW ENFORCEMENT MANAGEMENT AND ADMINISTRATIVE STATISTICS (LEMAS) 1, 4-5 (2013) [hereinafter LEMAS]. We utilized the 2013 data set, as that appeared to be the latest data set that was finalized and freely available at the time of our initial research. Since the purpose of the present measure is merely to illustrate the applications of the AF Method in the legal context, the age of the data should not be particularly relevant.

[56] A measure of this type would have value in that criminal justice-related literature has sought mechanisms for measuring multifaceted institutional performance. *See, e.g.*, Logan, *supra* note 2; Wright, *supra* note 1.

presented strictly as a means of reflecting how a measure based on the AF Method may be useful in the legal context.[57]

A. Description of the Data Set

The LEMAS survey sought to collect data from a nationally representative sampling of local and state law enforcement agencies located in the U.S.[58] A questionnaire was sent to 3,272 general purposes local and state law enforcement agencies, which included 895 sheriffs' offices, 2,327 local police departments, and the 50 primary U.S. state law enforcement agencies.[59] The questionnaire received an 86% response rate (with 2,822 of the 3,272 responding).[60] Accordingly, the final data set included information captured from such 2,822 responses, including from 717 sheriffs' offices, 2,059 local police departments, and 46 state agencies.[61] We suspect that, given the different nature of the three types of agencies, their capacities should be evaluated with reference to different criteria. Thus, for purposes of our measure, we focused exclusively on the local police departments, which we will refer to as "local law enforcement agencies" or simply "agencies."

B. Creation of Law Enforcement Capacity Index

In creating the LECI, we applied the steps for developing a bespoke multidimensional measure we set out earlier in this Article.[62] Consistent with Steps 1 and 2, we determined that the LECI's purpose was to measure the inadequacy of law enforcement capacity, and we established the local law enforcement agencies as the unit of identification. We then proceeded to select our indicators, assign them weights, and set relevant cutoffs, consistent with Steps 3 and 4.[63] These values are summarized in Table 3.

---

[57]   In this connection, if we had sought to create a measure for primarily substantive value, we might have made different choices, including in respect of the selected data, indicators, dimensions, weights and cutoffs, and analyses. We also would have included additional robustness checks.

[58]   LEMAS, *supra* note 55.

[59]   *Id.*

[60]   *Id.*

[61]   *Id.*

[62]   *See supra* Section II.A.

[63]   We have made such selections and determinations for illustrative purposes only, and with no representation that any are necessarily optimal or correct. Indeed, many of the indicators we have selected may implicate normative and policy issues, individual or entity preferences or standards, and/or political positions, and it is beyond the scope of this Article to weigh in on any of these issues. *See, e.g.*, JAMES MCCABE, AN ANALYSIS OF POLICE DEPARTMENT STAFFING: HOW MANY OFFICERS DO YOU REALLY NEED? (2012), https://icma.org/sites/default/files/305747_ Analysis%20of%20Police%20Department%20Staffing%20_%20McCabe.pdf (discussing police staffing); Cecelia Klingele et al., *Reimagining Criminal Justice*, 2010 WIS. L. REV. 953, 954-56 (2010) (same); Richard A. Bierschbach & Stephanos Bibas, *Rationing Criminal Justice*, 116 MICH.

We selected nine indicators drawn from responses to questions in the LEMAS survey, which we grouped into three dimensions of adequacy of agencies' capacities: "Policies and Procedures," "Staff and Training," and "Information and Technology." For convenience, we assigned an equal weight to each indicator, yielding an 11.11% weighting per indicator.

The Policies and Procedures dimension consisted of three indicators. The first indicator, "Mission Statement,"[64] identified as insufficient those agencies that did not have a written mission statement.[65] The second indicator, "Documentation on Use of Weapon,"[66] classified as insufficient those agencies that did not require any documentation after discharging a firearm or conducted energy device (e.g., a taser), or using a baton or OC spray/foam.[67] The third indicator, "Policy on Body Armor,"[68] considered insufficient those agencies that did not have a written policy on body armor.[69]

L. REV. 187, 232-37 (2017) (discussing resources); Seth Stoughton, *Law Enforcement's "Warrior" Problem*, 128 Harv. L. Rev. F. 225, 226-32 (2015) (discussing training); Yuri R. Linetsky, *What the Police Don't Know May Hurt Us: An Argument for Enhanced Legal Training of Police Officers*, 48 N.M. L. REV. 1 (2018) (same); Juleyka Lantigua-Williams, *How Much Can Better Training Do to Improve Policing?*, THE ATLANTIC (July 13, 2016), https://www.theatlantic.com/politics/archive/2016/07/police-training/490556/ (dis-cussing education, training, and recruitment); *Standards on Urban Police Function*, AM. BAR ASS'N (Dec. 5, 2018), https://www.americanbar.org/groups/criminal_justice/publications/criminal_justice_section_archive/crimjust_standards_urbanpolice/ (discussing pro-cedures, training and education, and recruitment for urban police); RICHARD SILBERGLITT ET AL., PRIORITY CRIMINAL JUSTICE NEEDS INITIATIVE, VISIONS OF LAW ENFORCEMENT TECHNOLOGY IN THE PERIOD 2024-2034 (2015), (discussing technology); Edwin Meese III & John Malcolm, *Introduction: "Criminal Justice at a Crossroads" Symposium*, 23 BERKELEY J. CRIM. L. 1 (2018) (same); Stephen E. Henderson, *A Few Criminal Justice Big Data Rules*, 15 OHIO ST. J. CRIM. L. 527 (2018) (same). Moreover, general limitations of empirical measurement also constrained our choices, in particular the quality and availability of relevant data. Finally, we eschewed considerations of indicator validity and reliability, which fall outside the scope of this methodological Article.

[64] Based on question E1 in the 2012 LEMAS survey. U.S. DEP'T OF JUSTICE, OFFICE OF JUSTICE PROGRAMS, BUREAU OF JUSTICE STATISTICS, LAW ENFORCEMENT MANAGEMENT AND ADMINISTRATIVE STATISTICS SURVEY 4 (2012), https://www.bjs.gov/content/pub/pdf/LEMAS_2013_CJ44_final_version.pdf [hereinafter LEMAS SURVEY]. Note that although the survey uses a 2012 date in its title, the LEMAS study used a 2013 date, since most questions in the survey involved a reference date in January 2013. LEMAS, *supra* note 55, at 188.

[65] LEMAS, *supra* note 55, at 188. By "written mission statement," the 2012 LEMAS survey means "[t]he agency's written statement of purpose that should guide the actions of the organization, spell out its overall goal, provide its general enforcement principles, and guide decision-making." LEMAS SURVEY, *supra* note 64, at B5.

[66] Based on question H2 in the 2012 LEMAS survey. LEMAS SURVEY, *supra* note 64, at 7.

[67] LEMAS, *supra* note 55, at 265-70. By "OC spray/foam," the 2012 LEMAS survey means "[a] chemical agent (a compound that irritates the eyes to cause tears, pain, and even temporary blindness) that is used in riot control, crowd control, and personal self-defense. Also known as pepper spray." LEMAS SURVEY, *supra* note 64, at B4. By "conducted energy device," the 2012 LEMAS survey means "[l]ess-lethal devices intended to deliver an electrical charge sufficient to momentarily disrupt a subject's central nervous system, enabling better officer control of the individual and causing minimal discomfort or injury." LEMAS SURVEY, *supra* note 64, at B1.

[68] Based on question H7 in the 2012 LEMAS survey. LEMAS SURVEY, *supra* note 64, at 7.

[69] LEMAS, *supra* note 55, at 283.

The Staff and Training dimension consisted of three indicators. The first indicator, "Number of Sworn Personnel,"[70] identified as insufficient those agencies that had less than one sworn person per 850 inhabitants covered.[71] The second indicator, "Minimum Education Requirement,"[72] considered insufficient those agencies in which the minimum educational requirement for new sworn hires was below the high school level or its equivalent (e.g., the GED).[73] The third indicator, "Community Policing Training,"[74] classified as insufficient any agency where less than half of the agency's full-time sworn personnel recruits, or full-time sworn personnel via in-service, received a minimum of eight hours of training in the previous twelve months on community policing issues.[75]

The Information and Technology dimension consisted of three indicators. The first indictor, "Website,"[76] considered insufficient any agencies without a website.[77] The second indicator, "Computerized Records,"[78] classified as insufficient agencies that did not maintain computerized records of incident-based statistics or officer narratives.[79] The third indicator, "Access to Information,"[80] identified as insufficient any agency in which the agency's patrol officers lacked direct electronic access to at least one of the following types of information: motor vehicle records, driver license records, criminal history, outstanding warrants, protection orders, or history at address.[81]

We set the overall inadequacy cutoff at one third. This means that agencies insufficient in at least one third of the weighted indicators are identified as inadequate in capacity. We set this cutoff because we assume that one or two insufficiencies are not enough to compromise the performance of the agencies.

---

[70]   Based on question A1 in the 2012 LEMAS survey and LEMAS's population information. *See* LEMAS Survey, *supra* note 64, at 1; LEMAS , *supra* note 55, at 12. The LEMAS survey defines "Sworn personnel" as "Law enforcement officers and deputies with general arrest powers." LEMAS Survey, *supra* note 64, at B4.

[71]   LEMAS, *supra* note 55, at 14-15.

[72]   Based on question C6 in the 2012 LEMAS survey. LEMAS Survey, *supra* note 64, at 3.

[73]   LEMAS, *supra* note 55, at 139-43.

[74]   Based on question E2 in the 2012 LEMAS survey. LEMAS Survey, *supra* note 64, at 4.

[75]   LEMAS, *supra* note 55, at 188-90. The 2012 LEMAS survey notes that "Community policing is a philosophy that promotes organizational strategies, which support the systematic use of partnerships and problem-solving techniques, to proactively address the immediate conditions that give rise to public safety issues, such as crime, social disorder, and fear of crime." LEMAS Survey, *supra* note 64, at B1.

[76]   Based on question F10 in the 2012 LEMAS survey. LEMAS Survey, *supra* note 64, at 6.

[77]   LEMAS, *supra* note 55, at 226-27.

[78]   Based on question F5 in the 2012 LEMAS survey. LEMAS Survey, *supra* note 64, at 5.

[79]   LEMAS, *supra* note 55, at 206-08.

[80]   Based on question F2 in the 2012 LEMAS survey. LEMAS Survey, *supra* note 64, at 5.

[81]   LEMAS, *supra* note 55, at 201-06.

Table 3: LECI - Dimensions, Indicators, Cutoffs, and Weights

| Dimension	Indicator	Insufficient if	Weight (%)
Policies and Procedures	Mission Statement	Does not have a written mission statement	11.11
	Documentation on Use of Weapon	No documentation is required after using firearm, baton, spray, or taser	11.11
	Policy on Body Armor	Does not have a written policy about body armor	11.11
Staff and Training	Number of Sworn Personnel	There is less than 1 sworn person per 850 inhabitants covered	11.11
	Minimum Education Requirement	Minimum education requirement for new sworn hires is below high school or equivalent	11.11
	Community Policing Training	Less than half of sworn recruits or sworn personnel received at least 8 hours of training during the previous 12 months on community policing issues	11.11
Information and Technology	Website	Does not have a website	11.11
	Computerized Records	Does not maintain computerized records of incident-based statistics or officer narratives	11.11
	Access to Information	Patrol officers have no direct electronic access to at least one of the following: motor vehicle records, driver license records, criminal history, outstanding warrants, protection orders, or history at address	11.11

C. Results and Analyses of Law Enforcement Capacity Index

Applying the LECI framework we set out in section III.B to the LEMAS data, we can gain both an overall picture of inadequacy and a more granular view of the components of inadequacy. We begin by estimating the overall incidence of inadequacy, intensity of inadequacy, and LECI. These three figures—along with their respective 95% confidence intervals—are presented in Table 4. The incidence shows that 38.1% of the agencies are inadequate. The intensity reflects that inadequate agencies are, on average, insufficient in 42% of indicators, which corresponds to almost four indicators. The LECI is 0.160, which means that the total insufficiencies experienced by inadequate agencies correspond to 16% of all possible insufficiencies. Such aggregated figures provide a broad summary of the agencies' inadequacy in capacity.

Table 4: LECI - Incidence, Intensity, and Inadequacy

| Index (k=33.33%)	Value	Confidence Interval (95%)	
Incidence (H, %)	38.1	35.2	41.0
Intensity (A, %)	42.0	40.6	43.4
Inadequacy or LECI	0.160	0.146	0.174

We then seek to identify the insufficiencies that drive inadequacy in order to better inform targeted corrective policies. To do this, we start by examining the indicators' censored headcount ratios, depicted in Figure 1 in the darker color. We use the censored headcount ratios because they focus only on the insufficiencies experienced by *inadequate* agencies, rather than the insufficiencies among *all* agencies. The insufficiencies among all agencies, or uncensored headcount ratios, are depicted in Figure 1 in the lighter color. To illustrate the importance of focusing on the censored headcount ratios, note that although 58.9% of the agencies are insufficient in Access to Information (uncensored headcount ratio), only 32.3% are inadequate *and* insufficient in Access to Information (censored headcount ratio). This means that 26.6% of the agencies are insufficient in Access to Information, but *not* inadequate. From the perspective of a policymaker, the insufficiency in Access to Information of these 26.6% agencies may be of less interest, since such insufficiency is not combined with enough additional insufficiencies such that the agencies' capacity might be compromised. The censored headcount ratios in Figure 1 reflect that the insufficiencies most prevalent among inadequate agencies are Access to Information, Community Policing Training, and Website. More than one quarter of the agencies are inadequate *and* insufficient in each of these indicators. On the other side of the spectrum, Minimum Education Requirement and Documentation on Use of Weapon do not appear particularly problematic, since less than 5% of the agencies are inadequate *and* insufficient in each of these indicators. Accordingly, a policymaker seeking to decrease inadequacy might consider targeting Access to Information, Community Policing Training, and Website, while taking no action on Minimum Education Requirement and Documentation on Use of Weapon.

Figure 1: Censored and Uncensored Headcount Ratios
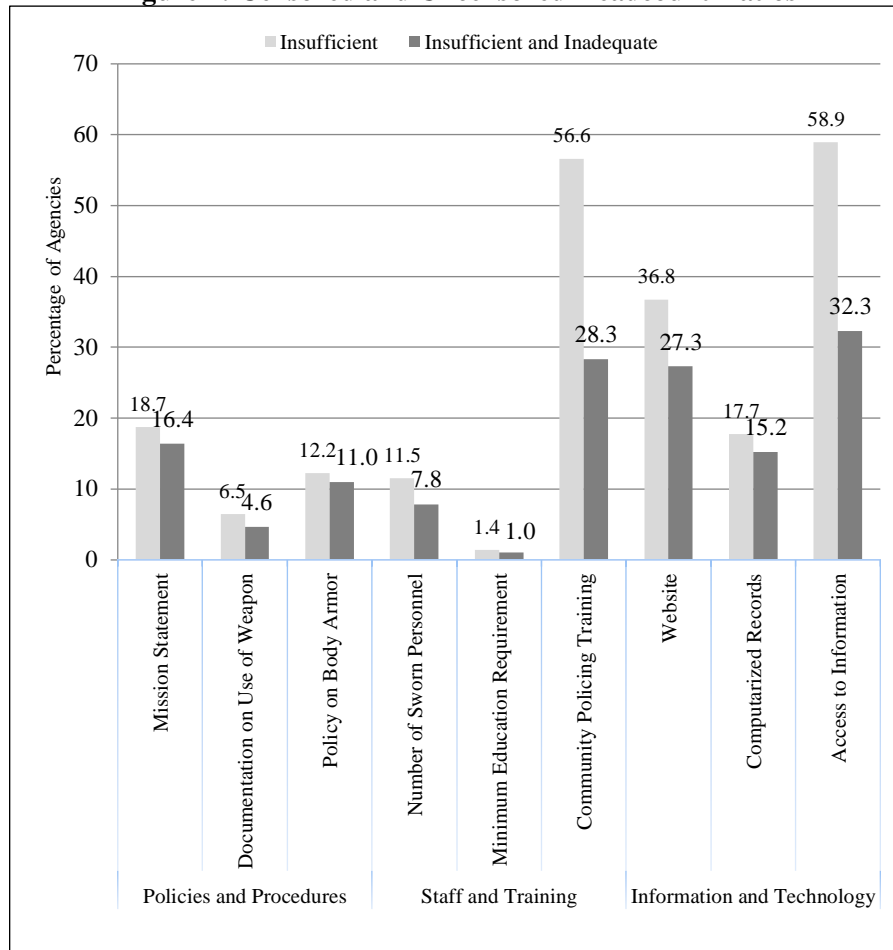
Another way we may seek to isolate the drivers of inadequacy is by quantifying the relative contribution of each indicator to the index (summarized in Figure 2).   Consistent with analysis of the censored headcount ratios, Figure 2 reflects that Access to Information, Community Policing Training, and Website make the largest relative contribution to the index, while Minimum Education Requirement and Documentation on Use of Weapon make the smallest.   At the dimension level, the dimension Information and Technology appears to be a key driver of inadequacy, accounting for more than half of the index, with insufficiencies in Access to Information alone accounting for 22.4% of inadequacy. The dimension Staff and Training has a relative contribution of 25.8%, primarily due to insufficiencies in Community Policing Training (19.7%). The dimension Policies and Procedures accounts for only 22.2% of the index, with

insufficiencies in Mission Statement representing 11.4% of the index. Calculating and presenting the relative contribution of each indicator complements the censored headcount ratio analysis and may further aid a policymaker in making targeted resource decisions.

Figure 2: Relative Contribution of Indicators
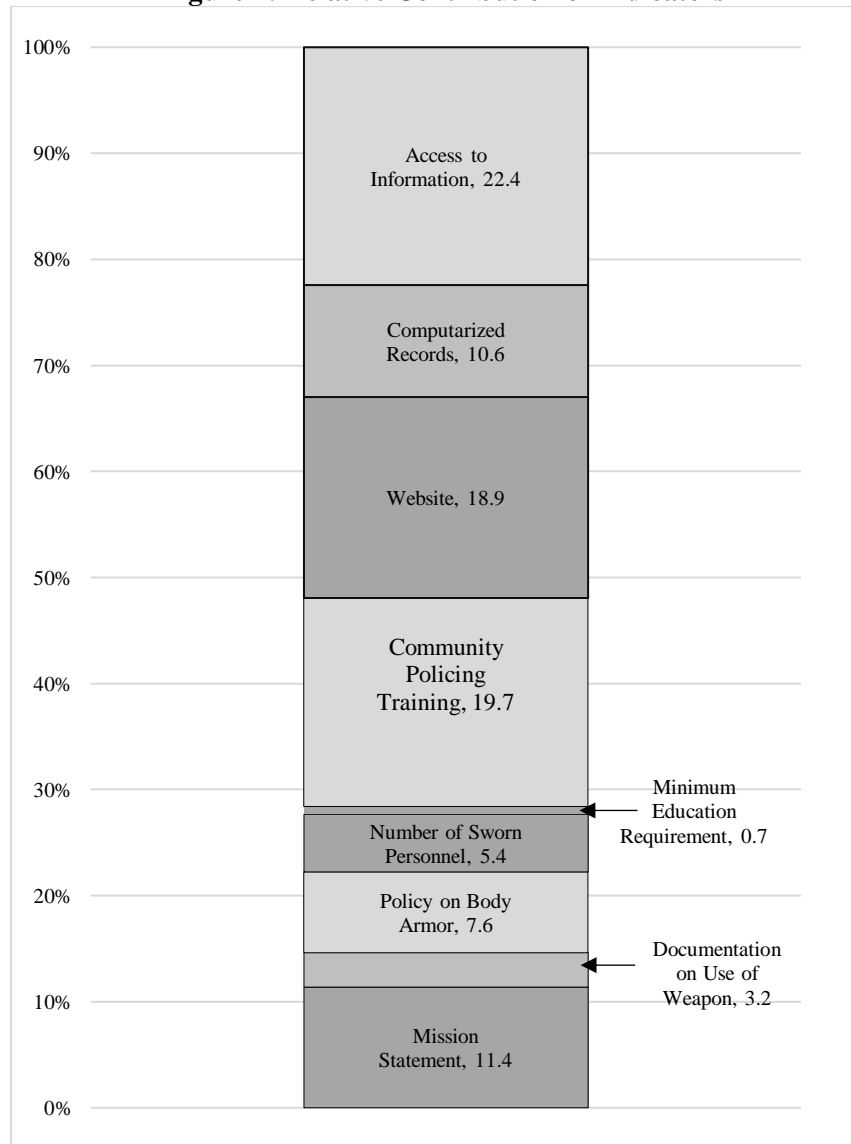
The calculations we have made so far have included analysis of *all* agencies studied across the country. Countrywide averages, however, may mask significant differences *across* groups of agencies. To address this, we

can use the framework to zoom in on the situation of particular groups. For instance, in Figures 3 and 4, we present the levels of inadequacy and composition of inadequacy of the agencies in the ten most populous states.[82] As Figure 3 shows, the percentage of agencies deemed inadequate varies widely, between 11.3% in Florida and 42.5% in Texas. Policymakers might, thus, wish to target attention toward states with higher percentages in Figure 3, such as Texas, Pennsylvania, and New York. Figure 4 shows that the main drivers of inadequacy vary across states.[83] For example, in California, the relative contribution of the dimension Policies and Procedures is almost null, and the main driver of inadequacy is the insufficiencies in the indicator Number of Sworn Personnel, with a relative contribution of 27%. In Texas, on the other hand, the dimension Policies and Procedures accounts for 34% of the inadequacy, and there seems to be practically no lack of sworn personnel, as the relative contribution of that indicator is 1%. This might mean that, in order to decrease state-level inadequacy, a policymaker in California might focus on Number of Sworn Personnel, while a policymaker in Texas might focus on Policies and Procedures.

Figure 3: Incidence of Inadequacy in Ten Most Populous States

---
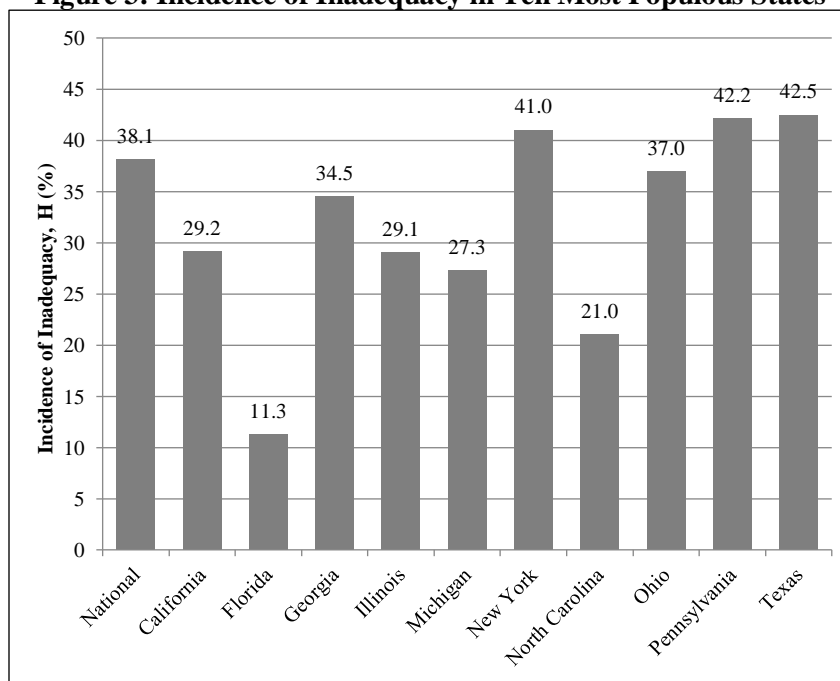
[82]   The purpose of this exercise is simply to illustrate the usefulness of the framework. Drawing statistically significant inferences regarding the agencies' inadequacies across states would require representative state-level data and an estimation of standard errors.

[83]   To facilitate readability, we have rounded numbers in this Figure to the nearest full percentage.

**Figure 4: Relative Contribution of Indicators
for Ten Most Populous States**
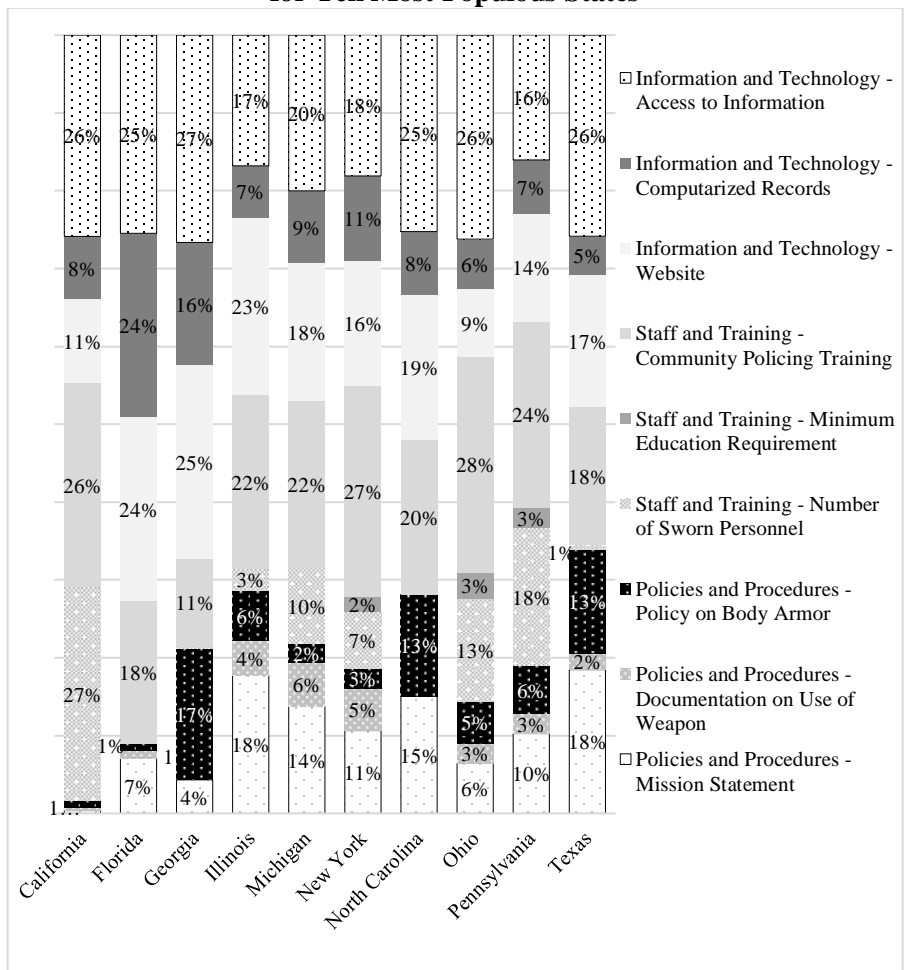
Since so many normative decisions are required when constructing a measure, it is important to analyze how sensitive the calculated results are to changes in the selected parameters. For example, it is helpful to examine the extent to which the ordering of the states by inadequacy depends on the normative inadequacy cutoff chosen. Figure 5 reflects the incidence of inadequacy for four selected states as a function of possible inadequacy cutoffs. The figure suggests that the incidence of inadequacy in Texas is greater than or equal to that in Florida regardless of the cutoff selected. This means the ordering of Texas and Florida would not be sensitive to the chosen cutoff. In contrast, the fact that the lines for Ohio and North Carolina intersect several times, reflects that comparisons between the incidence of

inadequacy in these two states is sensitive to the cutoff selected.[84] Where incidence of inadequacy is sensitive to the normative cutoff selected, one should review the results with this limitation in mind and consider calculating the measure with various cutoffs to obtain a more complete picture of inadequacy.

**Figure 5: Headcount Ratio for Selected States
at Different Inadequacy Cutoffs (k)**
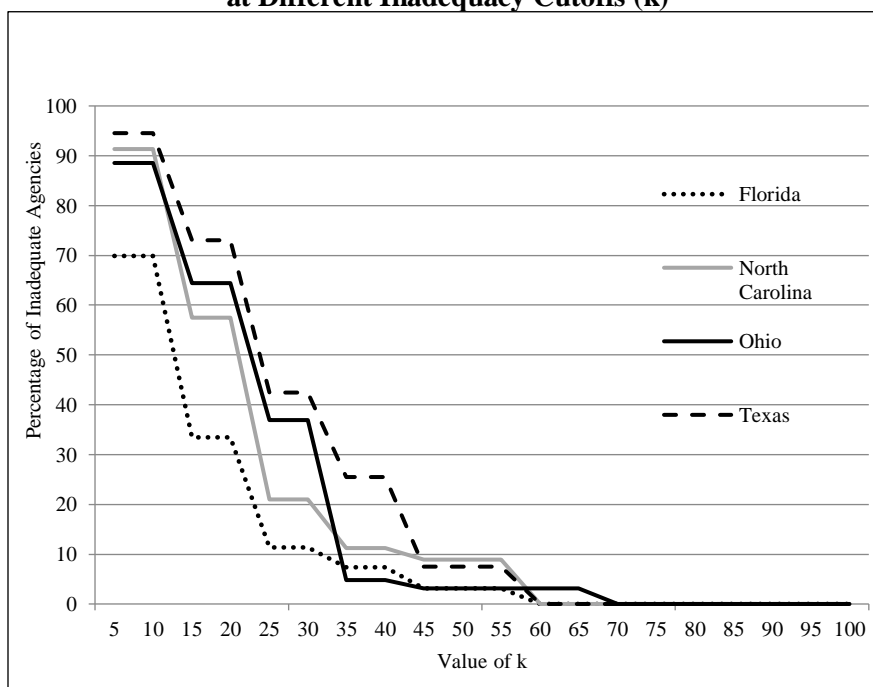
IV. CONCLUSION

The purpose of this Article was to illustrate the value of creating multidimensional measures in the legal context using the AF Method. Accurate measurement at the aggregate and granular levels drives good policy, and many legal phenomena worth measuring are of the type best suited to multidimensional measurement. It is hoped that our sample LECI has illustrated the ways in which legal researchers, professionals, and policymakers can use the framework to create measures fitted to their measurement needs. At a minimum, the legal community should conduct further research in the area of law and multidimensional measurement. In

---

[84]   Again, a rigorous analysis should take into consideration the standard errors.

particular, the accuracy and effectiveness of multidimensional measurement in the legal context would be greatly aided by the collection of better data, as well as research into the appropriate minimum standards that should be required to achieve justice.