

[*1] I. INTRODUCTION

Like all right-thinking law professors, I have tut-tut-tutted the law school rankings put out by the *U.S. News & World Report*. n1 [*2] Students accord it an objectivity it does not deserve! Ranking leads law schools to compete along the wrong dimensions! It creates incentives to cheat! I believed all of those things . . . and still do. But my recent work on democratic reform has led me to soften my views on rankings, perhaps even to adopt a contrarian view. It seems to me that law professors generally underestimate the case for rankings. Here, I will offer a necessarily brief argument for the other side. It is quite possible to concede every point made by the critics and still write, as I do here, in praise of rankings.

I should state up front that I now have a dog in this fight. I have proposed that Congress create a Democracy Index that ranks states and localities based on how well they run elections. In this article, I will ground my claims using this example. Resting my argument on this example is a bit of a cheat. A Democracy Index is easier to defend than, say, the *U.S. News & World Report* rankings, to which I will return at the end of this article. n2 First, the Democracy Index doesn't yet exist, so no one can (yet) accuse it of being badly designed. Second, some activities lend themselves more easily to measurement than others, and the data that would be included in a Democracy Index fall on the comfortable end of this spectrum. n3 Academics call election administration practices the "nuts and bolts" with good reason. These aren't the issues that have riven the elections community. n4 Nonetheless, the Democracy Index is designed to do what many rankings are designed to do: help the general public make sensible choices and put pressure on institutions to improve.

Part I briefly outlines the general arguments in favor of creating a Democracy Index. Part II identifies the three major costs associated with rankings. Part III outlines the often-overlooked benefits associated with rankings. Part IV concludes with a brief reflection on the debate over the utility of the *U.S. News & World Report* rankings.

[*3] II. THE DEMOCRACY INDEX

Before turning to the pros and cons of rankings, let me offer a bit of background on the Democracy Index. The Index would rank states and localities based on election performance. n5 It would turn on the basic questions that matter to voters: How long were the lines? How many ballots got discarded? How many machines broke down? n6 The goal of the Index is to provide a realistic, comparative baseline for evaluating how well an election system is working.

Although the proposal for the Index is quite concrete, the arguments for creating it are animated by a larger theme. We have a 'here to there' problem in election reform. n7 We have a good sense of the 'here' (the problems with our current system) and lots of ideas about the 'there' (how things ought to work in the future). But we need to focus more on the 'here to there'--how to create an environment in which meaningful reform might actually take root. n8 Election reform battles are usually fought on quite hostile terrain. Even a crisis like the 2000 presidential debacle prompted only modest change in how we administer elections. For that reason, we should focus on small-scale interventions that will make bigger, better reform possible. 'Here to there' proposals are not meant to be the journey's end; they are designed to smooth the path that leads there.

If we want to create a more receptive environment for election reform, two immediate priorities spring to mind. First, we should focus on changing the terms of a debate that, thus far, has not resulted in a great deal of movement on reform. One reason that election reform is hard to pass is that the key players in the reform process--voters, judges, the media, and policymakers--almost always lack a yardstick for judging reform debates. Reform debates tend either to operate in the range of vague generalities or to involve nitty-gritty questions that demand a high level of expertise to resolve. One of our goals, then, should be to create [*4] better yardsticks for judging ongoing debates. Ideally, these yardsticks should work not just for elites who are capable of absorbing complex debates, but for voters who will, at best, rely on shortcuts and heuristics to referee fights over election reform.

Second, we should focus on the mechanisms that spur policy diffusion. There is robust social science literature on policymaking mimicry. That literature suggests that shortcuts play a useful role in diffusing best practices. n9 Policymakers and bureaucrats rely on [*5] shortcuts all the time. In choosing what policy to adopt, they look to the policies adopted by similar institutions or to widely held professional norms. Few people have the time to work through all of the normative and practical implications of every decision. It thus makes sense to look to one's neighbors or to the practices of one's peer group as a source of distilled wisdom. n10

Professional peer pressure is also involved in policy diffusion. n11 If you know that your legislative counterparts or professional peers have adopted a particular policy, you will feel pressure to do so for reasons that often have little to do with the policy itself. Professional peer pressure, in other words, works as well as its high school variants. If our goal is to get from 'here to there,' the question is whether we can create other institutional and decision-making shortcuts that would build on these insights to speed the spread of good policy.

The Index is just such a shortcut. It gives voters, policymakers, and bureaucrats a shorthand for making better decisions. n12 Voters, for instance, would have a means of evaluating the performance of election officials--something that should create at least modest incentives for politicians to pay attention to performance. The Index would also give voters a means of refereeing debates between reformers and election officials by providing them with a yardstick to evaluate the debates. The Index might even help generate support for reform among voters because it would give voters something to have an [*6] opinion about. While voters do not have views on the arcane policies, they do have a view on whether their state should be doing worse than its neighbors.

The Index should be similarly helpful for policymakers. It gives them a baseline--an information shortcut for refereeing debates between the election administrators who work for them and the reformers who lobby them. n13 While top policymakers may be reluctant to hold election officials accountable based on the necessarily atmospheric opinions of reformers, they are likely to be convinced by hard numbers and comparative data.

A ranking provides a useful shorthand in a second way: it helps flag policymaking priorities, thereby enabling top-level officials to distinguish between a modest glitch and a serious problem. In the long run, the performance data generated by the Index can help policymakers identify the innovation needle

in the haystack of widely varying practices.

Finally, the Democracy Index might help build professional norms among election administrators by jumpstarting a conversation about best practices. When we think about improving a system, we generally assume that the pressure for reform comes from the outside. But the long-term health of any system depends largely on administrators policing themselves based on shared professional norms. Indeed, professional norms may ultimately be more important to a well-run system than pressures from the outside. By providing a focal point for election administrators' attention and a shortcut for identifying best practices, the Democracy Index might help disseminate the types of professional norms on which a well-functioning bureaucracy depends.

III. A FEW CONCESSIONS

Any proposal like the Democracy Index inevitably runs into a serious set of questions, most of which have to do with the problems generically associated with rankings. There is little question that rankings have costs. Rankings create at least three kinds of problems:

[*7] . People imbue the rankings with an objectivity they don't deserve.

. They may induce institutions to compete along the wrong dimensions.

. They create an incentive to cheat. n14

The first problem is the natural consequences of distillation. Ranking requires a trade-off between precision and accessibility, and there are costs associated with this trade-off no matter what choice you make. The second two issues are what you might call 'happiness problems'; they occur if a ranking starts to get traction, but they can undermine its success in the long term.

A. *The Consequences of Distillation*

Rankings simplify. It is an inevitable consequence of trying "to provide *one* answer to a question when that answer depends on several bits of data," in the words of Oxford's Stein Ringen. n15 Distilling information can serve many useful ends, but any effort to rank necessarily involves a trade-off between precision and accessibility or "rigor and intuition." n16

Objectivity. Many worry about rankings because people think they are precise and objective, and they aren't. Attaching a number to an assessment lends it an aura of absolute truth. People are sure that the institution ranked first is better than the one ranked second, and they think that a ranking conveys meaningful information about the distance between, say, first and second or thirty-fourth [*8] thirty-fourth and thirty-fifth. Indices suggest not only precision, but another quality associated with mathematics: objectivity.

Needless to say, the reality is quite different from the perception. Ranking requires a large number of discretionary (and thus debatable) choices. Every stop along the way--deciding what to measure, how to measure it, and how to add the measurements together--requires a normative judgment.

David Roodman, chief architect of the Commitment to Development Index (CDI), is well aware of these trade-offs. Roodman's unusual background in math and communications gives him an acute sense of ranking's theoretical shortcomings and its practical utility as a "communications vehicle." n17 The challenge involved in ranking, he jokes, is "to do something that is analytically

impossible in a way that is analytically credible." n18

Just to ground the argument a bit, let me focus on what's likely to be the most debatable choice for any ranking: how to weight the data. Even when most people can agree on what to measure and how to measure it, there will be considerable disagreement about how to aggregate the data into a single ranking.

The trade-off between precision and accessibility is particularly acute in this context because the most sophisticated weighting techniques are likely to be the least transparent. The easiest, most transparent strategy is to pick commonsense categories and average the scores assigned to each category, just as one would calculate a grade point average. I propose just such a strategy for the Democracy Index. Needless to say, there will be other ways to aggregate the data (after all, what are the odds that everything is equally important?).

Why would one choose an equal weighting strategy, as I did and as many index designers have done? The reason is simple: other weighting strategies are just as debatable and a good deal less transparent. As Ringen observes, "If the weights are not [*9] objectively known, . . . go[ing] by the simplest assumption [is a sensible choice]." n19

While I think the decision to weight categories equally is defensible, no one would suggest that there is an easy way to resolve these debates, save perhaps the clueless husband in the *New Yorker* cartoon who asks his wife, "You want a child, I want a dog. Can't we compromise?" n20 As a practical matter, the only way to settle these debates is to settle them.

B. Happiness Problems

In addition to the two problems noted above--both having to do with the trade-off between precision and accessibility--there are at least two other potential pitfalls involved with ranking. Both involve competition run amok: a ranking can encourage institutions (1) to compete along the wrong dimensions and (2) to cook the books. Each is thus some variant of a happiness problem. Competition is, of course, exactly one hopes an index would promote, so both of these developments would be a heartening sign that one's index has gotten some traction. But competition can have perverse consequences if the index is poorly designed.

Competing Along the Wrong Dimensions. Rankings are designed to spur healthy competition, but they can sometimes cause people to compete along the wrong dimensions. When a poorly designed index starts to get traction, it can lead institutions to do unproductive, even silly things, to improve their standing.

Academics are especially quick to identify this problem because they have long lived with the *U.S. News & World Report* rankings, which are infamous for causing foolishness of all sorts. Say the word "ranking" to a law professor, and he or she will immediately remind you of all the ridiculous ploys that schools have used to improve their standing. Columbia Law School, for instance, pushed its faculty to take their leaves in the spring rather than the fall because student-teacher ratios are assessed only in the fall. As a result, the school had to hire thirty-two part-time [*10] teachers to accommodate spring teaching needs. n21 In order to jack up its score on student expenditures, the University of Illinois' law school counted the fair market value of its students' Westlaw/Lexis subscriptions (which totaled up to a cool \$ 8.78 million). Given that both research services heavily discount their fees in order to woo

future users, that "expenditure" was eighty times what Illinois actually paid. n22 Stanford Law School's entrepreneurial dean, Larry Kramer, has devoted part of his deanship to convincing the central university to let the law school "write a check" for its utilities rather than have the university deduct them automatically from student tuition. The reason for this accounting switch? It would allow the law school to count these expenses as student expenditures. n23 "The notion that I'm losing students because of this is insane," Kramer told the *New York Times*. n24

If the Democracy Index were poorly designed, it could lead to as much silliness as the *U.S. News & World Report* rankings, creating problems that are more serious than the accounting hocus-pocus described above. Take fraud. Most voters care about fraud, so it would be perfectly sensible to include a fraud metric in the Index. The question, however, is how to measure fraud without creating perverse incentives. We don't want unsubstantiated fraud prosecutions or roving posses of state officials accosting voters outside of polling places. These techniques have long been associated with vote suppression, and one would hardly want to give partisan officials an excuse to use them.

The Index might also create more systemic problems. Dan Esty, who helped put together the Environmental Performance Index (EPI), insists that "we measure what matters." n25 We can't always measure everything that matters, however. Data-driven analysis creates a risk that people will neglect important issues that [*11] can't be captured in a statistic. n26 Some data will be too costly to gather; some performance dimensions will be too difficult to quantify. A ranking might lead states to compete on the things that can be measured while ignoring those that can't. Imagine, for instance, that it is too difficult to assess whether a registration system is easy for voters to navigate. States might be reluctant to put money into building a better registration system when they can improve their score, even if only marginally, by investing in something else.

Cheating. A final potential cost associated with ranking is cheating. The worry is that institutions will cook the books to improve their rankings. Like the concern about states competing along the wrong dimensions, this worry is a variant of the happiness problem. For example, if the Democracy Index were having such a powerful effect on election officials that they were tempted to cheat, we would already have come a long way. Nonetheless, for an index to be a trustworthy guide, the underlying data must be dependable.

IV. IN PRAISE OF RANKINGS

It is possible to make all of the concessions above and still write in praise of rankings. That is not to say that rankings are always good or that any particular ranking is worth the candle. But there is a tendency among the professoriate to condemn rankings categorically, simply because the one we know best is flawed. n27 That is a mistake. We don't think wheels are a bad idea simply because we've had a flat. While I would not put rankings up there with the invention of the wheel, rankings are pretty useful. As I explain below, however, the utility of rankings is best understood when we ask a simple question, "As opposed to what?"

[*12] *A. We Often Simply, and with Good Reason*

There is no question that rankings simplify. The question is whether that is a cause for celebration or regret. Academics tend to assume the latter. Indeed,

they sometimes use the term 'simplify' as if it were synonymous with 'oversimplify,' as if any type of shorthand is presumptively illegitimate.

That is a mistake. We often simplify. Policymaking--or decision making of any sort n28--would be impossible without shortcuts. If all shorthand were eliminated, we wouldn't have a GDP and thus couldn't distinguish between an economic blip and a recession. Congress would never stop holding hearings because there would always be more testimony to collect. Consumer Reports would go out of business.

Even *disaggregated* data are a form of shorthand. As Dan Esty notes, "Quantification is about distillation." n29 The raw ingredients of the Democracy Index, for example, are stand-ins for a vast and complicated process that no individual could possibly evaluate firsthand. The very purpose of data is to distinguish between what Esty calls "signal" and "noise." n30 Consider, for instance, Roger Angell's evocative description of baseball box scores:

[A] box score is more than a capsule archive. It is a precisely etched miniature of the sport itself . . . [that] permits the baseball fan, aided by experience and memory, to extract from a box score the same joy, the same hallucinatory reality, that prickles the scalp of a musician when he glances at a score of Don Giovanni and actually hears basses and sopranos, woodwinds and violins. n31

[*13] Because shorthand is inevitable in a decision-making process, the real question is what kind of shorthand to use. For instance, when I've presented my proposal for a Democracy Index, people sometimes bristle at the idea of voters using shorthand to evaluate the way our elections are run. The question often asked of me is this: Why not present voters with a full range of information rather than 'spoon feed' them a ranking?

Though the concern is well taken, it misstates the question. The choice is not between spoon feeding voters or providing them with a full buffet of information. Voters will inevitably use some sort of shorthand in casting a ballot. The question is what kind of shorthand to supply.

Consider, for instance, how most voters cast their ballots. They usually know very little about the substantive positions of the candidates they elect. n32 Yet voters make surprisingly good decisions about how to cast a vote. Their decisions are by no means perfect and reveal predictable biases. But voters have figured out a pretty good strategy for choosing a candidate without sorting through the huge amount of information relevant to that decision.

How do voters do it? They use the party label as a shorthand--what political scientists would term a "heuristic"--in choosing a candidate. n33 The label *Democrat* or *Republican* functions like a Good Housekeeping Seal of Approval. It tells the voter that the candidate in question subscribes to values or policy preferences that are close enough to the voter's to choose him or her. n34 As several scholars have explained, if a voter "knows the [*14] big thing about the parties, he does not need to know all the little things." n35

Political scientists have devoted a lot of energy to making party cues function more effectively for a simple reason: they are a good deal better than the other types of shorthand that voters might use. n36 Without the party heuristic, voters would be more likely to base their votes on something unappetizing, such as a candidate's race or gender. Or they might cast ballots

randomly so that voter preferences are disconnected from electoral outcomes. The basic defense of party labels is not that they are perfect--far from it--but that they are the best thing we've got. If you ask a political scientist whether it is a good idea for voters to rely on party cues, the likely response will be a sarcastic, "As opposed to what?"

As Opposed to What? If we think about that question in the context of election administration, a ranking system looks a good deal more appealing. Think about the proxies voters are likely to use today in casting their vote for election officials. The best bets seem to be (1) anecdotal evidence, (2) news about a widely reported crisis, or (3) partisan cues. For all its potential shortcomings, a ranking system is superior to each of these alternatives.

Anecdotal evidence is, of course, just that. While a bewildering number of academics think that what their taxi driver said on the drive to the conference constitutes reputable proof, a glitch here and there is not good evidence of a full-fledged problem. A ranking system, in contrast, focuses voters on the [*15] bigger picture, directing their attention to systemic concerns instead of the modest anomalies that can afflict even well-run systems. It also directs their attention to the good as well as the bad and the ugly, revealing which states and localities have done an especially impressive job of running elections.

Even evidence of a crisis may not be a useful guide for voters. While the worst-run systems are more vulnerable to a crisis, not all badly run systems will experience one. Indeed, given the dearth of the data, we cannot definitively rule out the possibility that recent brouhahas have happened in relatively well-run systems--places that just happened to be in the path of a turnout tsunami. Crisis-based voting also has the flavor of closing the barn door after the horse has been stolen. Voters need a tool that will help them prevent crises, not just react to them.

Finally, partisan cues don't provide a dependable heuristic for voters in this context. A party label can tell a voter whether a candidate is liberal or conservative--something that may map on to particular approaches to campaign finance or felon disenfranchisement. But in choosing an election administrator, voters need shorthand for evaluating professionalism and performance, and the party cue does not help. Democrats and Republicans are equally susceptible to running elections badly.

For all of these reasons, the Democracy Index has the potential to provide voters with a much-needed shorthand for casting a vote. By conveying information about the 'big thing' in election administration--a rough sense of how well the system performs overall--it enables voters to make sensible decisions without knowing all of 'the little things' buried in the data.

One might think it is too easy to make the case for rankings for voters, who are relatively unsophisticated actors in the reform process. But are rankings defensible for decision makers with more knowledge and power, like policymakers? The answer is yes. In many ways, the Index serves the same purpose for top-level policymakers as it does for voters: it gives them a baseline for refereeing debates between the election administrators who work for them and the reformers who lobby them. Policymakers hear plenty of untrustworthy arguments from administrators who aren't doing their job properly. Many grow tired of the insistent drum beat for change emanating from the reform community. Top-level

[*16] policymakers have to pick sides, and they do not have time to work through all the details. They need an information shortcut to guide them.

While top policymakers may be reluctant to hold election officials accountable based on the necessarily atmospheric opinions of reformers, they are likely to be convinced by hard numbers and comparative data. Election administrators can talk all they want about what they have done. But they cannot get around the stark reality of a ranking: Is the system working or not? And why is the state next door doing so much better?

A ranking provides a useful shorthand in another way: it helps flag policymaking priorities. Legislators and governors are often bombarded with information. They hear lots of complaints, listen to lots of requests for funding, and sift through lots of reports. What they need is something that helps them separate the genuine problems from run-of-the-mill complaints--a means of distinguishing the signal from the static. A ranking can perform that role as it focuses on systemic problems and provides a realistic baseline for judging performance.

Consider, for instance, what occurred in Mexico when the first version of the EPI (then called the Environmental Sustainability Index) was released. n37 Environmentalists had spent a lot of time trying to convince Mexico it had a problem. They ended up spending most of their time addressing low-level bureaucrats. When the first version of the EPI came out, ranking Mexico in the bottom fifth of the countries evaluated, it caught the attention of Mexico's President. Dan Esty's team received dozens of calls and e-mails from Mexican officials up and down the political hierarchy, all complaining about Mexico's ranking and, eventually, trying to figure out how to fix it. Mexican bureaucrats cared because the President cared.

B. The Objectivity Problem

None of this is meant to suggest that rankings are objective--they are not--but neither are the other shortcuts on [*17] which people usually rely. The key, then, is to make a judgment call as to whether the trade-off between precision and accessibility--a trade-off present for most decision-making shortcuts--is a sensible one.

For instance, let us return to the weighting issue, which I discussed above. n38 The way a ranking weights categories is inevitably invoked in any fight over a ranking. Weighting is necessary in order to simplify, and it is necessarily subjective. Nonetheless, we should not overstate the costs of simplification. Stein Ringen offers the most cogent argument for ranking. "We *always* rely on conventions of some kind or other in any effort at measurement, and indexing is in that respect not extraordinary," he writes. n39 "As long as we use sensible conventions and explain the procedures, there is nothing unscientific in it." n40

Dan Esty would surely agree. The designers of the EPI, after consulting numerous experts, chose to divide the environmental performance metrics they were using into two main categories--measures of environmental health (which includes things like child mortality rates, drinking water quality, and sanitation) and measures of ecosystem vitality (which includes things like timber harvest rates, overfishing, and renewable energy). n41 Remarkably, although this fifty-fifty weighting is in theory the most contestable part of the EPI, in practice it's been the least contested. n42 Esty attributes that

fact to two things. The fifty-fifty breakdown is intuitive. And no one has come up with a better strategy.

Similarly, David Roodman, the chief architect of the CDI, argues that in the absence of a consensus on weighting, an equal weighting strategy represents the simplest and most transparent method for assembling the CDI. Roodman told me that months of consulting with academic experts convinced him that his simple weighting strategy was the right one. Why? No one else had a [*18] suggestion that could garner a consensus. For every category, there was always one expert or another suggesting it was entitled to more weight than it got. n43 For this reason, Roodman chose the most transparent one. "Weighting things equally," he says, "says that I don't know the answer." n44

Moreover, there are ways to discourage people from vesting a ranking with too much objectivity. The people who put out an index should be careful about overselling its objectivity or accuracy. Problems in the data should be discussed forthrightly. The designers of an index should note its lack of precision. They should also identify--and explain--the judgment calls they made along the way. All of the data should be publicly available, and the mechanism for aggregating the data should be as transparent as possible. Ideally, the materials accompanying an index should not only make the designers' choices clear, but should also show the ways in which those choices affect the ranking itself. For instance, the website for the Democracy Index could allow people to rerun the rankings themselves using their own criteria n45 --the election geek's version of 'choose your own adventure.'

Nonetheless, these mitigating strategies can not come close to eliminating the costs described above. Rankings work in part because people do not look past the number. But even if you think that people put too much faith in an index, the more important question is whether they would otherwise put their faith in something sillier. The fact that there is not an 'objective' shortcut for people to use does not mean that all shortcuts are created equal.

We're back to the question, "As opposed to what?" A ranking will surely oversimplify the state of affairs. But, as Ringen observes, "While some information gets lost, something else is gained." n46 Take the Democracy Index as an example. Reams of comparative data cannot give us a clear view of how jurisdictions are performing overall. As with party labels, rankings tell voters about the 'big thing' even if they lose track of the 'little things.' A well-designed index fares particularly well when it is compared to [*19] the other shorthand that citizens use in evaluating voting processes--anecdote, haphazard evidence of a crisis, or partisan labels. The public places unwarranted faith in each of these heuristics. Each leads to oversimplification and mistake of a more significant sort than a well-designed index will. And not one of them gets us any closer to improving our failing system. In this context, something seems better than nothing. n47

The bottom line here depends almost entirely on what you think a ranking is for. n48 If the goal is simply to convey information, the answer is obvious: don't rank. Presenting data in disaggregated form will almost always be better than ranking. But if the goal is to improve the policy-making process--to correct a failure in the political market--the only thing that beats a good ranking is a better one.

V. HAPPINESS PROBLEMS

A. Competing Along the Wrong Dimensions

As noted above, a badly designed ranking can cause people to compete along the wrong dimensions, and virtually any ranking will at least reduce the incentive to compete along dimensions that cannot be reduced to a measurement. These are serious problems. But here, again, the question to ask is what does a world without rankings look like?

To begin, the quality of the ranking matters a great deal. There are badly designed rankings, and there are well-designed ones. A [*20] badly designed ranking will likely produce foolish competition. A well-designed ranking should produce at least some healthy competition.

There are also concrete steps one can take to produce a well-designed ranking. The first is regular reevaluation and revision of the index. A second strategy for avoiding foolish competition is to create a comprehensive index. A sparse index has its virtues, but there are costs associated with parsimony. If there is only a handful of metrics in the index, it is easier to improve one's standing by focusing on one or two. Further, if lots of performance information is left out of the index, institutions can easily divert resources from the parts of its system that aren't being measured to the parts that are.

Consider, for instance, how this applies to a Democracy Index. If a state tried to increase its fraud score by engaging in techniques that deterred voters from casting a ballot, it might lower its score on any "ease of voting" metrics included in the Index. Comprehensiveness might similarly help with the problem of resource diversion. The more that is measured, the fewer tasks the state can neglect with impunity. For instance, returning to the example above, if the state neglects its registration system, it may find that lots of voters--mistakenly thinking that they've properly registered--will show up to vote. That would create administrative headaches for poll workers and longer lines for properly registered voters--problems that would, in turn, reduce the state's overall score.

Here again, the costs and benefits associated with ranking are flip sides of the same coin. An index encourages election officials to compete along the dimensions it measures. That means one needs to consider two issues in deciding if this is a good thing. First, the quality of the ranking must be considered. If the ranking is well designed, 'teaching to the test' is all to the good. It is a problem, however, if the test is poorly designed.

Second, we need to know what kind of teaching takes place when there's no test. Right now, we are in a world with no test; we lack even the most basic data for evaluating the performance of our election system. If the Democracy Index works, it will surely reorient state and local priorities, perhaps causing them to neglect concerns that the Index doesn't measure. The cost might be [*21] significant enough to eschew data-driven analysis if most of the basic components of election administration can't be captured in a statistic. We're back to the question, "As opposed to what?" A well-designed Democracy Index is surely better than the alternative . . . a world without data--one with no test at all.

B. Cheating

A final potential cost associated with ranking is cheating. Like the concern about states competing along the wrong dimensions, this worry is a variant of

the happiness problem. For a ranking to be useful, the underlying data must be dependable.

There are two obvious strategies for dealing with the problem of cheating. The first is to rely on data from outside sources whenever possible. In the context of election administration, voter surveys, for instance, can provide a pretty good mechanism for gathering basic data on many parts of the election process. "Testers" can similarly help us evaluate information that is otherwise in the state's control.

For the pieces of data that can come only from the state, the obvious solution to cheating is verification. For example, the designers of a Democracy Index might use random sampling to double-check state disclosures. Random sampling might be prohibitively expensive on a large scale. But it can be used to 'spot-check' state data from time to time. In spot-checking state disclosures, designers of the Index might even be able to piggyback on existing research. Political scientists spend a good deal of time using random samples to investigate basic questions about how the election system works, and the designers of the Index could use that research as an outside check on internal state reporting.

The designers of the Democracy Index could also follow the lead of other index architects, who rely on many sources to verify information passed on by state officials. The designers of the Government Performance Project, for instance, use a triangulation strategy--asking the same question of many different actors in the state. n49 Similarly, the architects of the Democracy Index might [*22] talk to local polling officials, civil-rights watchdogs, and local reporters to identify problems that have gone unreported by state officials. This sort of qualitative read should help designers of the Index figure out whether they are working with decent quantitative information.

If state legislatures or Congress decide to mandate that states disclose performance data, they could also create backstops against cheating. Congress might, for instance, require states to certify the data or obtain an outside expert's blessings, just as corporations are required to do under the Sarbanes-Oxley Act. n50

Finally, designers of the Index could use the ranking system to punish states for faking the data. The people who put together the EPI, for instance, routinely toss data that don't seem plausible. n51 A state that is caught cooking the books could be punished by imputing the lowest possible number for the relevant portion of the ranking. Or the Democracy Index could include a 'disclosure' component that would reward states that adopt the sort of certification practices described above.

Cheating may be the most difficult problem posed by ranking. It is hard to detect and thus costly to avoid. While cheating would, in some ways, be a sign of the ranking's success--no one would bother to cheat if the ranking didn't matter--it would also jeopardize the ranking's power in the long run. The "As opposed to what?" question is tougher here, too. It is hard to argue that it's better to have institutions rampantly cheating to improve their rankings than not to have a ranking system at all.

There are a few reasons not to throw in the towel, however. Cheating is most likely to happen when a ranking is getting traction--when it is starting to affect debates and influence behavior. And the more traction a ranking gets, the

more reasons its supporters will have to invest in it, perhaps providing the resources necessary to create alternative sources of data or better policing strategies. Further, the more comprehensive the ranking [*23] becomes, the harder it will be to fake enough data to affect it. If the only metrics on which an institution performs well are those that involve self-reporting, people will suspect that something is afoot. To put it differently, as a ranking grows in importance, its designers should have more tools available to police the cheating that might accompany that growth.

VI. CONCLUSION

If we take a hard look at the costs associated with ranking, it is clear that they are genuine and genuinely important. The question is whether, in light of these costs, the game is still worth the candle.

We cannot answer that question by evaluating these trade-offs in the abstract. We have to think about how they play out in the real world, and that means answering another question, "As opposed to what?" In the abstract, the costs seem quite weighty: rankings are accorded more objectivity than they deserve, they can encourage institutions to compete along the wrong dimensions, and they create an incentive for cheating. But the *real-world* alternative may be worse. In the elections context, for instance, it is one in which voters rest their decisions on far sillier shorthand, localities feel no pressure to compete, and there isn't any test to cheat on.

This brings me back to the *U.S. News & World Report* rankings. They are powerful, and they are imperfect. There is good reason to demand that they improve or to try to create a better ranking. Nonetheless, it is a mistake to use the *U.S. News & World Report* ranking as an excuse to demand an end to ranking. Even an admittedly flawed ranking, like this one, has its merits. Ask yourself, for instance, whether a world without the *U.S. News* is really as attractive as some make it out to be. It is not hard to imagine college students basing their choice on far sillier criteria--general reputation (something that would lead many to apply to the nonexistent Princeton Law School), where Felicity went to school, n52 or what Mom and Dad think.

[*24] Editorial writer Meg Greenfield once observed that "Everybody [is] for democracy--in principle. It's only in practice that the thing gives rise to stiff objections." n53 It's just the reverse for rankings. It's easy to be against rankings in principle. It's only in practice that they start to look good.

FOOTNOTES:

n1 For the substantive pros and cons of the *U.S. News & World Report* rankings, with most people firmly in the con camp, compare Michael Sauder & Ryon Lancaster, *Do Rankings Matter? The Effects of U.S. News & World Report Rankings on the Admissions Process of Law Schools*, 40 LAW & SOC'Y REV. 105, 106 (2006), and Ronald G. Ehrenberg, *Method or Madness? Inside the U.S. News & World Report College Rankings*, J.C. ADMISSION, Fall 2005, at 29, and Shuky Ehrenberg, A Question of Rank 10 (Nov. 2007) (unpublished Edgar M. Cullen prize paper, Yale Law School), available at http://digitalcommons.law.yale.edu/ylsspps_papers/18/, and STEPHEN P. KLEIN & LAURA HAMILTON, ASS'N OF AM. LAW SCHS., THE VALIDITY OF

THE U.S. NEWS AND WORLD REPORT RANKING OF ABA LAW SCHOOLS (1998), <http://www.aals.org/reports/validity.html>, with Mitchell Berger, *Why the U.S. News and World Report Law School Rankings are Both Useful and Important*, 51 J. LEGAL EDUC. 487 (2001), and Symposium, *The Next Generation of Law School Rankings*, 81 IND. L.J. 1 (2006). It is worth noting that some critics of the *U.S. News & World Report* rankings quarrel with its methodology, but not with the idea of ranking itself. For instance, Brian Leiter, one of the most trenchant critics, believes that a good ranking is possible. See Brian Leiter, *Welcome to Brian Leiter's Law School Rankings*, <http://www.leiterrankings.com/> (last visited Dec. 14, 2009).

n2 See *infra* pp. 23-24.

n3 HEATHER K. GERKEN, *THE DEMOCRACY INDEX: WHY OUR ELECTION SYSTEM IS FAILING AND HOW TO FIX IT* 52 (2009).

n4 *Id.*

n5 GERKEN, *supra* note 3, at 5.

n6 *Id.* at 53.

n7 *Id.* at 6.

n8 *Id.*

n9 This literature is vast, so the proffered studies are only a sampling drawn primarily from the work of sociologists and political scientists. See, e.g., HENRY R. GLICK, *THE RIGHT TO DIE: POLICY INNOVATION AND ITS CONSEQUENCES* 44-46, 203 (1992); ANDREW KARCH, *DEMOCRATIC LABORATORIES: POLICY DIFFUSION AMONG THE AMERICAN STATES* 3 (2007); KAREN MOSSBERGER, *THE POLITICS OF IDEAS AND THE SPREAD OF ENTERPRISE ZONES* 102, 108, 111 (2000); RICHARD ROSE, *LESSON-DRAWING IN PUBLIC POLICY: A GUIDE TO LEARNING ACROSS TIME AND SPACE* 5-6, 31-33, 136-37 (1993);

Steven J. Balla, *Interstate Professional Associations and the Diffusion of Policy Innovations*, 29 AM. POL. RES. 221, 221-22, 229-33, 235-40 (2001); Frances Stokes Berry & William D. Berry, *Innovation and Diffusion Models in Policy Research*, in THEORIES OF THE POLICY PROCESS 169, 171-72, 175-76 (Paul A. Sabatier ed., 1999); Frances Stokes Berry & William D. Berry, *State Lottery Adoptions as Policy Innovations: An Event History Analysis*, 84 AM. POL. SCI. REV. 395, 395-97 (1990); Frances Stokes Berry & William D. Berry, *Tax Innovation in the States: Capitalizing on Political Opportunity*, 36 AM. J. POL. SCI. 715, 715, 721-22 (1992); Virginia Gray, *Innovation in the States: A Diffusion Study*, 67 AM. POL. SCI. REV. 1174, 1174-76, 1181-82 (1973); Michael Mintrom, *Policy Entrepreneurs and the Diffusion of Innovation*, 41 AM. J. POL. SCI. 738, 738-41, 753-54 (1997); Michael Mintrom & Sandra Vergari, *Policy Networks and Innovation Diffusion: The Case of State Education Reforms*, 60 J. POL. 126, 136-38 (1998); Anne Schneider & Helen Ingram, *Systematically Pinching Ideas: A Comparative Approach to Policy Design*, 8 J. PUB. POL'Y 61, 65-66 (1988); Jack L. Walker, *The Diffusion of Innovations Among the American States*, 63 AM. POL. SCI. REV. 880 (1969) (offering a comprehensive discussion of diffusion theory); David L. Weimer, *The Current State of Design Craft: Borrowing, Tinkering, and Problem Solving*, 53 PUB. ADMIN. REV. 110, 115-17 (1993); Harold Wolman & Ed Page, *Policy Transfer Among Local Governments: An Information-Theory Approach*, 15 GOVERNANCE: AN INT'L J. POL'Y, ADMIN., & INSTITUTIONS 477, 493, 497 (2002). For a critical take on some of this work, see generally Christopher Z. Mooney, *Modeling Regional Effects on State Policy Diffusion*, 54 POL. RES. Q. 103 (2001) (questioning whether regional effects on diffusion are as pronounced as prior work has suggested). For an examination of interstate diffusion in different countries, see Katerina Linos, *How Can International Organizations Shape National Welfare States?: Evidence from Compliance with European Union Directives*, 40 COMP. POL. STUD. 547, 562-63 (2007); James M. Lutz, Note, *Emulation and Policy Adoptions in the Canadian Provinces*, 22 CANADIAN J. POL. SCI. 147, 147-49 (1989); Dale H. Poel, *The Diffusion of Legislation Among the Canadian Provinces: A Statistical Analysis*, 9 CANADIAN J. POL. SCI. 605 (1976) (presenting empirical data on diffusion among the Canadian provinces); Note, *When Do Policy Innovations Spread?: Lessons for Advocates of Lesson-Drawing*, 119 HARV. L. REV. 1467 (2006) (discussing diffusion generally and offering critique of many prior works). For a survey of the literature on diffusion among nonstate organizations, see generally David Strang & Sarah A. Soule, *Diffiision in Organizations and Social Movements: From Hybrid Com to Poison Pills*, 24 ANN. REV. SOC. 265 (1998) (reviewing scholarly arguments on diffusion within nonstate organizations and social movements).

n10 GERKEN, *supra* note 3, at 86.

n11 *Id.* at 86-88.

n12 For an in-depth exploration of these ideas, see *id.* at chs. 3-4.

n13 GERKEN, *supra* note 3, at chs. 3-4.

n14 A fourth problem associated with some rankings is the problem of targeting. When large institutions (like states) are ranked, they are treated as if they are unitary actors. But in most large institutions, many decision makers influence the activity being evaluated. As a result, an unhappy consumer or voter may put pressure on someone who doesn't have much control over the results. Imagine, for instance, that the Democracy Index ranked State X forty-seventh in the nation, and voters began to pester the Secretary of State. The Secretary of State might legitimately complain that the real problems with the system stem from decisions made at the local level or might blame the legislature or vice versa. How do voters figure out who's at fault?

n15 STEIN RINGEN, *WHAT DEMOCRACY IS FOR: ON FREEDOM AND MORAL GOVERNMENT* 283 (2007).

n16 Interview with Dan Esty, Professor, Yale Univ. (Oct. 24, 2007).

n17 Interview with David Roodman, Ctr. for Global Dev. (Feb. 1, 2008).

n18 David Roodman, Ctr. for Global Dev., *Building and Running an Effective Policy Index: Lessons from the Commitment to Development Index*, Mar. 13, 2006, at 2, <http://www.cgdev.org/content/publications/detail/6661> (click "Download" hyperlink).

n19 KINGEN, *supra* note 15, at 284.

n20 Michael Maslin, Cartoon, *NEW YORKER*, July 23, 2001, at 38.

n21 Alex Wellen, *The \$ 8.78 Million Maneuver*, *N.Y. TIMES*, July 31, 2005, § 4A, at 18.

n22 *Id.*

n23 *Id.*

n24 *Id.*

n25 Environmental Performance Index: Home, <http://epi.yale.edu/Home> (last visited Dec. 14, 2009).

n26 For a useful analysis of this problem, see DANIEL C. ESTY & REECE RUSHING, CTR. FOR AM. PROGRESS, GOVERNING BY THE NUMBERS: THE PROMISE OF DATA-DRIVEN POLICYMAKING IN THE INFORMATION AGE 38-39 (2007), http://www.americanprogress.org/issues/2007/04/pdf/data_driven_policy_report.pdf.

n27 *But cf.* Leiter, *supra* note 1. Brian Leiter is a noteworthy exception to this observation. *Id.*

n28 Behavioral economists have built an entire field around people's tendency to rely on heuristics and the costs and benefits associated with that strategy. For a seminal work in this vast field, see Amos Tversky & Daniel Kahneman, *Judgment Under Uncertainty: Heuristics and Biases*, 185 SCI. 1124 (1974).

n29 Interview with Dan Esty, *supra* note 16.

n30 *Id.*

n31 ROGER ANGELL, *ONCE MORE AROUND THE PARK: A BASEBALL READER* 4 (1991). Many thanks to Doug Chapin for sharing this with me.

n32 See SAMUEL L. POPKIN, *THE REASONING VOTER: COMMUNICATION AND PERSUASION IN PRESIDENTIAL CAMPAIGNS* (1991).

n33 For early work in what is now a vast literature, see, e.g., JOHN H. ALDRICH, *WHY PARTIES? THE ORIGIN AND TRANSFORMATION OF POLITICAL PARTIES IN AMERICA* (1995); BERNARD R. BERELSON, PAUL F. LAZARFELD & WILLIAM N. MCPHEE, *VOTING: A STUDY OF OPINION FORMATION IN A PRESIDENTIAL CAMPAIGN* (1954); ANGUS CAMPBELL, PHILIP E. CONVERSE, WARREN E. MILLER & DONALD E. STOKES, *THE AMERICAN VOTER* (1960); V.O. Key, Jr. & Frank Munger, *Social Determinism and Electoral Decision: The Case of Indiana*, in *AMERICAN VOTING BEHAVIOR* 281, 286-87 (Eugene Burdick & Arthur J. Brodbeck eds., 1959).

n34 See, e.g., Richard R. Lau & David P. Redlawsk, *Voting Correctly*, 91 *AM. POL. SCI. REV.* 585, 590 (1997); Arthur Lupia, *Dumber than Chimps? An Assessment of Direct Democracy Voters*, in *DANGEROUS DEMOCRACY? THE BATTLE OVER BALLOT INITIATIVES IN AMERICA* 66, 69 (Larry Sabato, Howard R. Ernst & Bruce A. Larson eds., 2001); Arthur Lupia, *Shortcuts Versus Encyclopedias: Information and Voting Behavior in California Insurance Reform Elections*, 88 *AM. POL. SCI. REV.* 63, 63 (1994); POPKIN, *supra* note 32.

n35 BERELSON, LAZARFELD & MCPHEE, *supra* note 33, at 321.

n36 Indeed, a major movement within political science insists that we need strong, cohesive parties in order to give voters a better predictive cue as to how candidates will vote. Better party cues, the argument goes, means greater accountability. This notion of "responsible party government" was first endorsed by the American Political Science Association's Committee on Political Parties in 1950. See *Am. Political Sci. Ass'n, Toward a More Responsible Two-Party System: A Report of the Committee on Political Parties*, 44 *AM. POL. SCI. REV.* 1 (Supp. 1950).

n37 See generally *Environmental Performance Index: Contents*, <http://epi.yale.edu/Contents> (last visited Dec. 14, 2009) (providing the rankings and more information on the EPI).

n38 *See supra* p. 8-9.

n39 RINGEN, *supra* note 15, at 284.

n40 *Id.*

n41 YALE CTR. FOR ENVTL. LAW & POLICY, PILOT 2006 ENVIRONMENTAL PERFORMANCE INDEX 285 (2006), http://www.yale.edu/epi/2006EPI_Report_Full.pdf.

n42 Interview with Dan Esty, *supra* note 16.

n43 Interview with David Roodman, *supra* note 17.

n44 *Id.*

n45 Thanks to Christine Jolls for offering this suggestion.

n46 RINGEN, *supra* note 15, at 284.

n47 None of this is to say that we will improve the political process by giving people bad information. As America's Manolo-shod philosopher, Carrie Bradshaw, once observed, at some point "the art of compromise become[s] compromising." Something is not always better than nothing. But a pretty good something usually is.

n48 Oxford's Stein Ringen frames the question similarly but offers a different answer. He argues that the decision whether to rank depends on whether one wants to "guide decision makers towards good decisions" or "assist towards a rational political process." RINGEN, *supra* note 15, at 285. In his view, these categories boil down to "measurement" and "social reporting." *Id.* at 286. When decision makers just need information about facts on the ground in order to decide, indexing is a sensible strategy. But, argues Ringen, when the issue at stake involves "political, normative, or moral questions," indexing improperly short-circuits the deliberative process. *Id.* at 285-86.

n49 Interview with Philip Joyce, Professor, George Washington Univ. (Sept. 19, 2007).

n50 15 U.S.C. §§ 7201-7214 (2006). It is worth noting that this requirement has generated a good deal of debate and controversy among corporations and academics. Congress would have to think hard about the lessons learned from Sarbanes-Oxley before taking this path.

n51 Interview with Dan Esty, *supra* note 16.

n52 Bill Lucia, *Applying to College Isn't Easy*, TIMES UNION (Albany, N.Y.), Dec. 2, 1999, at D1. It is tough to figure out whether the decision of the eponymous character in the television show, *Felicity*, to attend New York University (NYU) led to an uptick in NYU's applications. Media reports of the claim are ubiquitous, but they are generally accompanied by precious little proof and, occasionally, by denials from NYU public relations folks (one of whom insisted that the application bump was due to "recent movement away from the traditional 'Jeffersonian model' of education"). *Id.*

n53 Meg Greenfield, Op-Ed., *The People's Revenge*, WASH. POST, June 14, 1978, at A27.