Comparing energy efficiency scoreboard methodologies and results

Wolfgang Eichhammer
Fraunhofer Institute for Systems and Innovation Research ISI
Breslauer Str. 48
DE-76139 Karlsruhe, Germany
and Utrecht University
Heidelberglaan 2
NL-3584 CS Utrecht, The Netherlands
wolfgang.eichhammer@isi.fraunhofer.de
w.a.eichhammer@uu.nl

Shruti Vaidyanathan
American Council for an Energy-Efficient Economy ACEEE
529 14th Street NW, Suite 600
Washington, DC 20045
USA
svaidyanathan@aceee.org

Meegan Kelly
American Council for an Energy-Efficient Economy ACEEE
529 14th Street NW, Suite 600
Washington, DC 20045
USA
mkelly@aceee.org

Barbara Schlomann
Fraunhofer Institute for Systems and Innovation Research ISI
Breslauer Str. 48
DE-76139 Karlsruhe
Germany
barbara.schlomann@isi.fraunhofer.de

Keywords

data, energy efficiency policy, policy evaluation, methodology, evaluation methods, scoreboard, scorecard

Abstract

We carry out in this paper a comparison of the methodologies and results of Energy Efficiency Scoreboards. Several such scoreboards have been developed in recent years notably the International Energy Efficiency Scorecard and the State Energy Efficiency Scorecard by ACEEE and the ODYSSEE-MURE Energy Efficiency Scoreboard. We also provide an overview of other policy scoreboards in the energy field and look to lessons learned from other fields beyond energy (such as scoreboards to measure innovation or educational capabilities of countries). We analyse the methodologies used to compose the scoreboard, the possible impact of the methodologies on results, the data sources and data requirements for establishing the scoreboards, the ranking procedures, and the results. We also discuss how the scoreboards have impacted or may impact policy making.

Introduction

Scoring of efforts and achievements is present in everyone's life when it comes to evaluating the performance of individuals. We personally experience scoring and evaluation through marks during school and university or during the professional life. We may also be involved in the evaluation of firms, by comparing a firm's performance to another in the same field, or in the evaluation of countries, with respect to innovation capability, educational performance or competitive positioning in trade, for example.

Scoring and ranking is based on two important psychological components of human nature. First, we all seek the comparison to others to develop benchmarks for our own performance. We ask ourselves, "Am I good or bad compared to the performance of others?" Second, when we observe our performance is bad compared to others, we may again ask ourselves, "How are they achieving better performance? Can I learn from them and achieve similar or even better results?"

From this we derive two basic objectives for the scoring and ranking procedure we generally call a *scoreboard*: (1) measuring performance, and (2) comparing entities. The first step implies that a scoreboard defines a number of criteria according to which the performance of participants is gauged. The second step involves ranking which helps to compare the entities with each other. These two components are the main reason why various forms of scoring are widespread in many domains.

Energy efficiency was not included in comprehensive scoreboard efforts for many years, with the notable exception of energy efficiency indicators, which have been used to study energy-use and compare countries since the eighties. Full scoreboards that include measurements of both energy efficiency performance and policies have only been developed in recent years. Therefore, there is still a lot to gain by comparing the methodologies that have been developed in those efforts. It is also a timely moment to discuss what role energy efficiency scoreboards have had or may have in future policy making.

International scoreboards for energy efficiency

Several full scoreboards have been developed in recent years. Notably, the ODYSSEE-MURE Energy Efficiency Scoreboard and the American Council for an Energy-Efficient Economy (ACEEE) State and International Energy Efficiency Scorecards. The following section presents the ODYSSEE-MURE and ACEEE scoreboards more extensively and briefly introduces other scoreboards that include an energy efficiency component.

THE ODYSSEE-MURE ENERGY EFFICIENCY SCOREBOARD

The ODYSSEE-MURE Energy Efficiency Scoreboard has recently been developed and published (ODYSSEE-MURE Scoreboard, 2015) under the long-running European ODYSSEE-MURE project on energy efficiency indicators and policies, which has provided large analytical inputs for the development of European energy efficiency policies. Very recently, the project has been providing inputs for the impact assessment of the proposal for the amendment of the Energy Efficiency Directive (EED) aiming to increase the European energy efficiency target to 30 % for 2030 (European Commission, 2016). The objective of the ODYSSEE-MURE Scoreboard for Energy Efficiency Indicators and Policies is to assess and score for all EU countries in a common European scoreboard: (i) energy efficiency level (present state of energy efficiency), (ii) energy efficiency progress (trends in energy efficiency) and (iii) energy efficiency policies (future potential for energy efficiency progress).

At present, the ODYSSEE Scoreboard for Indicators combines the first two issues in one scoreboard; the MURE Scoreboard for Energy Efficiency Policies scores separately energy efficiency policies. In the ongoing project, the ODYSSEE-MURE Scoreboard for Energy Efficiency Indicators and Policies shall combine all three issues in one single scoreboard with an equal weighting of the three components.

THE ACEEE STATE AND INTERNATIONAL ENERGY EFFICIENCY SCORECARDS

Table 1 lists all of ACEEE's scorecards. For simplicity, this paper focuses on two, the state and international scorecards.

ACEEE published its first comprehensive approach to scoring and ranking US energy efficiency policies with its *State Energy Efficiency Scorecard* in 2007 (Eldridge et al. 2007). The Scorecard was developed as a means to document best practices, recognize leadership, and encourage federal action to catch up with and complement state efforts. A set of metrics and a scoring system identified where and how the best progress on energy efficiency was being achieved. The report's state ranking and review of policy approaches also offered guidance to policy makers looking for practical models to follow.

Since the first edition, ACEEE has continued to update the State Scorecard, providing an annual picture of both progress and decline in state energy efficiency policy. Governors, legislators, regulators, and citizens look to the report as a source for reliable information, which led to the development of additional energy efficiency scorecards, including an *International Energy Efficiency Scorecard*, a *City Energy Efficiency Scorecard*, and a *Utility Energy Efficiency Scorecard*.

The International Scorecard is published bi-annually, with three editions since 2012. It measures national level energy efficiency by comparing a selection of countries across a set of common metrics. The results provide insights into the best policies and practices employed around the world and a global benchmark nations can use to paint a picture of how efficient they are. This report has an additional goal of placing the US on a global stage to build awareness among federal policy makers of world trends and identify opportunities to learn from leading nations on energy efficiency.

OTHER ENERGY EFFICIENCY SCOREBOARDS

Next to these three scoreboards there exist other energy efficiency scoreboards or scoreboards with an important energy efficiency component. Ranking in these scoreboards is a critical step and participants at the lower end tend to feel unfairly treated and may criticize the methodology used. This has led to approaches where the entities are not fully ranked but the criteria for scoring are shown separately. Depending on how the performance is aggregated across the different criteria, one could establish a scoreboard based on:

- *A* "strong ranking principle": Weights are established for the different criteria and the "score" across the different criteria is added to allow establishing a ranked list. The underlying methodology for the ranking is generally a more or less transparent multi-criteria analysis.
- A "medium strong ranking principle": The overall result for each criterion is shown in the form of spider or flower graphs. No overall ranking occurs across all criteria.
- A "weak ranking principle": the performance of each criterion is merely reported without any summary view.

In each case, a decision must be made on the set of criteria (metrics) and – in the first two cases – the scores for the metrics. Table 2[1] gives an overview of other scoreboards in the field of energy. It identifies whether a strong, medium, or weak ranking principle is used and notes what entities are included.

SCOREBOARDS IN FIELDS OTHER THAN ENERGY EFFICIENCY

Scoreboards have been developed in many fields of application, including innovation, science and technology, education, company performance and resource efficiency. In fact, these scoreboards are much better implemented and well-known than energy efficiency scoreboards at present. What we can learn from these other scoreboards is that scoring has tremendous impact on the discussion of policy success, both with respect to inputs to policy as well as outputs. Not all of these scoreboards include country rankings; some international organisations are reluctant to provide marks to countries, but are often inconsistent with their approaches. For example, the OECD Science, Technology and Industry Scoreboard does not provide a ranking, while the OECD PISA ranking – in fact one of the most influential scoreboards in history – does. Table 3 gives an overview of scoreboards in fields other than energy.

1. IEA country scorecards specifically related to combined heat and power: Some of the more recent scorecards use a "five star" ranking principle to benchmark countries against global practice.

Table 1. ACEEE Scorecards.

Ranking publication	Jurisdiction	Frequency
State Energy Efficiency Scorecard (http://aceee.org/state-policy/scorecard)	51 US states and 3 territories	Annually since 2006
International Energy Efficiency Scorecard (http://aceee.org/portal/national-policy/international-scorecard)	23 countries	Bi-Annually since 2012
City Energy Efficiency Scorecard (http://aceee.org/local-policy/city-scorecard)	51 US cities	Bi-Annually since 2013
Utility Energy Efficiency Scorecard	50+ largest utility providers	Forthcoming in 2017

Table 2. Overview of scoreboards and ranking principles in the field of energy.

Type of Scoreboards	Ranking Principle	Notes
ACEEE International Energy Efficiency Scorecard	strong	23 countries worldwide
ODYSSEE-MURE Energy Efficiency Scoreboard	strong	28 EU countries (+Norway, Switzerland, Serbia)
ARAB Future Energy Efficiency Index AFEX (http://www.rcreee.org/projects/arab-future-energy-index™-afex)	strong	17 countries in Arab region
CO_2 Scorecard (http://www.co2scorecard.org/)	medium	Large number of countries worldwide
Energy Efficiency Watch http://www.energy-efficiency-watch.org/	medium	28 EU countries
IEA Scoreboard 2011 (https://www.iea.org/publications/freepublications/publication/IEA_Scoreboard2011.pdf)	weak	IEA countries
IEA country scorecards specifically related to combined heat and power https://www.iea.org/chp/countryscorecards/	weak/medium	IEA countries

Table 3. Overview of scoreboards in fields other than energy.

Type of Scoreboards
OECD Science, Technology and Industry Scoreboard: The scoreboard measures how science, technology and innovation foster competitiveness, productivity and growth. http://www.oecd.org/sti/oecd-science-technology-and-industry-scoreboard-20725345.htm
OECD PISA: PISA is an international student assessment that has enormous impact on national education systems. Countries such as Germany and Brazil improved their student performance and made their education systems more inclusive as a result of PISA benchmarks. http://www.oecd.org/pisa/
European Innovation Scoreboard: The European Innovation Scoreboard – previously Innovation Union Scoreboard – provides a comparative analysis of innovation performance in EU Member States, other European countries, and regional neighbours. It assesses relative strengths and weaknesses of national innovation systems and helps countries identify areas they need to address. https://ec.europa.eu/growth/industry/innovation/facts-figures/scoreboards_de
Telekom Innovation Indicator: The indicator shows Germany's performance in the field of innovation capabilities with respect to other countries. www.innovationsindikator.de/
Shanghai Academic Ranking of World Universities: The ranking presents the world's Top 500 universities annually based on transparent methodology and third-party data. http://www.shanghairanking.com/
Balanced Scorecard at company level: Balanced Scorecard is used extensively in business and industry, government, and non-profit organizations worldwide. It provides a management tool to help align business activities to the vision and strategy of the organization. BSCs have evolved from an early use as a simple performance measurement framework to a full strategic planning and management system. The "new" balanced scorecard transforms an organization's strategic plan from an attractive but passive document into the "marching orders" for the organization on a daily basis. It provides a framework that not only provides performance measurements, but helps planners identify what should be done and measured. It enables executives to execute their strategies. https://www.balancedscorecard.org/Resources/About-the-Balanced-Scorecard
European Resource Efficiency Scoreboard: The Resource efficiency scoreboard is a tool and user interface for presenting key indicators relating to natural resources. For this scoreboard, a limited set of already available indicators was selected, covering as many as possible of the themes and subthemes identified in the Roadmap to a Resource Efficient Europe. It is a three tier system based on a lead indicator, a dashboard of indicators and a set of theme specific indicators http://ec.europa.eu/eurostat/web/environmental-data-centre-on-natural-resources/resource-efficiency-indicators/resource-efficiency-scoreboard

Energy efficiency scoreboards and policy making

Energy efficiency scoreboards have had an effect on policy-making in multiple ways. At the highest level, these reports and rankings create a sense of healthy competition amongst the evaluated entities with regards to progress on energy efficiency. This sense of competition generates discussion with key stakeholders and decision makers to eventually identify policy weakness and steer the conversation towards more ambitious energy reduction agendas. ACEEE has an experience of ten years when it comes to the State Scorecard, and five years for the International Scorecard. Other scoreboards outside the energy field have similar experience. We therefore mainly draw on these experiences when it comes to the impact and relevance of scoreboards on policy making.

For example, the Governor of Mississippi, the lowest ranking state in ACEEE's 2012 *State Energy Efficiency Scorecard,* publicly expressed the need to improve its standing from last place and used rankings to motivate the inclusion of energy efficiency in the state's energy plan and the passage of efficiency policy in the legislature (Bryant, 2012). The following year, Mississippi became the most improved state in the nation, jumping from 51st to 47th in the 2013 *State Energy Efficiency Scorecard* (Office of Commissioner Brandon Presley, 2013). In another example, the City of Los Angeles, the second-most populous city in the United States enumerated its goal to "Be a Top 10 city as rated by the American Council for an Energy Efficient Economy (ACEEE)" in its 2015 sustainability plan (PLAN, 2015).

Energy efficiency scoreboards can also be used to justify implementation of more ambitious policies and methodologies may even guide specific policy development. For instance, the state of Pennsylvania cites the *State Energy Efficiency Scorecard* in its 2015 Climate Change Action Plan, comparing it's score to the top-rated state and listing recommendations of a suite of policies that help reduce energy consumption a and move the state toward becoming a national leader (Pennsylvania, 2016). Another state, New Hampshire, cites the criteria used to define an energy efficiency resource standard (EERS) in ACEEE's Scorecard as a basis for its rules establishing a state energy savings target (New Hampshire, 2016). In the long run, scorecards could serve as road maps for future energy and climate plans, by providing lagging countries, states, and cities that are interested in address the energy efficiency of their economies with a detailed list of best practice policies in each end use energy sector to follow.

Similarly, the ODYSSEE-MURE Scoreboard has been presented to the larger public in Germany, in cooperation with the German Company Initiative Energy Efficiency DENEFF, with large resonance in the media and the policy sphere (DENEFF/ Fraunhofer ISI, 2015). While the experience with the ODYSSEE-MURE Scoreboard is still recent, it demonstrates strong potential to have an impact.

Finally, ACEEE's suite of scorecards engage energy offices and other stakeholders with regards to data collection and therefore help to identify data gaps in publicly available data. This then encourages state, local, and federal governments to make this data more easily accessible and sometimes goes as far as prompting them to collect the relevant data from scratch. Data collection is a critical step in empowering entities to manage their own energy use and scorecards can help facilitate positive changes.

Scoreboards outside the energy field, e.g. in the field of education as well as innovation capabilities, have shown that scoreboarding can largely structure the debate when it comes to the evaluation of public policies and their comparison across jurisdictions. Particularly striking is the impact of the PISA scoreboard on education in Germany. The organisation Learning-is-Open (2013) states: "When the first PISA tests in 2000 placed German students well below the average in OECD countries for reading and literacy, the nation was shocked. The revelation sparked a nationwide debate about Germany's school system and what was needed to improve it. Published in 2001, the PISA 2000 results revealed that the German system was not providing equal opportunities for all. Students from disadvantaged backgrounds were particularly at risk. Rather than basing the choice of secondary school solely on student achievement in elementary school, the system was actually denying opportunity to disadvantaged students". The same source to conclude: "Thanks to the combination of reforms and a nationwide effort to raise performance, Germany's education outcomes are improving. In the PISA 2009 tests, Germany's lowest-achieving students did better than in 2000, while its highest-achieving students maintained 2000 levels and the negative impact of students' socio-economic background diminished. From 21st place in 2000, Germany rose to equal 15th place alongside Sweden, just ahead of Ireland and France but behind the U.S."

This shows the potential of such scoreboards to stir debate and policy progress and the value they present in the energy field where scoring is a relatively new exercise.

Scoreboard methodologies

In this section, we analyse the ODYSSEE-MURE and ACEEE scoreboards in more detail. We document the methodologies used to compose the scoreboard, discuss possible impacts of the methodologies on results, identify the data sources and data requirements for establishing the scoreboards, and review the ranking procedures.

DATA REQUIREMENTS AND METHODOLOGY OF THE ODYSSEE-MURE ENERGY EFFICIENCY SCOREBOARD

ODYSSEE-MURE developed so far separately the ODYSSEE Scoreboard based on energy efficiency indicators and the MURE scoreboard for energy efficiency policies. In the future, both scoreboards will be combined.

ODYSSEE indicator scoreboard

Methodology

The *ODYSSEE Indicator Scoreboard* scores the level and progress of countries in energy efficiency, by sector (households, transport, industry and services). The scoring is done for a list of selected indicators representative of end-uses, transport mode or sub-sector. The scoring methodology is based on the OECD Composite Indicators methodology, which gives normalized scores across the countries within a range of 0-1, according to the formula in the inset (minimum and maximum are defined by the 3 lowest and highest values)[2].

2. Direction: -1 if a decrease in the indicator is favoured and +1, if an increase is favoured).

$$\frac{\text{Indicator} - \text{min indicator}}{(\text{max indicator} - \text{min indicator}) * \text{direction}} + 0.5*(1 - \text{direction})$$

Scoring is done separately for four sectors (households, transport, industry and services):

- Buildings (4 end-uses[3])/Transport (3 + 2 modes[4]): several indicators (each combined 50 % by level and 50 % for the trend). The score of each indicator is multiplied by the weight of the indicator (for details on the weighting see ODYSSEE-MURE methodological documents); which has been defined on the basis of the share of the end-use or transport mode in the sector consumption. The normalized indicator scores multiplied by their weight have been added to obtain the score of each sector.
- Industry: energy intensity at EU average industry structure, and one composite indicator summarizing energy efficiency trends (ODEX). For ODEX weighting occurs with the share in energy consumption.
- Service sector: distinguishes fuel and electricity consumption per employee (and also level/trend).

Ranking procedure

To obtain the global score, the score of each sector is multiplied by the weight of sectors, which are defined according to the actual sector shares in total final energy consumption of each country, and then added together.

The ranking occurs through a strong ranking principle, i.e. country rankings are shown. However, there was debate among the participants of the ODYSSEE-MURE project whether the worse countries are to be shown. This is the reason, why presently only the ten best countries out of 30 are shown. Another idea discussed was to show beyond the tenth position the countries only by bands, i.e. 11–15, 16–20 etc.

Data collection and requirements

The data for the indicator scoreboard are collected through the regular data collection process of the ODYSSEE-MURE project, which is also used for other purposes (e.g. analysis for the EU Commission in the frame of impact assessments, for the establishment of national reports on energy efficiency progress, for country or sector profiles etc.). Thus the scoreboard does not require specific additional data. This regular data collection differentiates the Indicator Scoreboard from other scoreboards where the data first need to be compiled. In the ODYSSEE-MURE project data are collected at national level but through a harmonised template with harmonised definitions of data. Data are then checked by the technical coordination of the projects, which assures that the final data set is harmonised across the EU countries.

3. Heating, other thermal uses, appliances, penetration solar water heaters.

4. Cars, trucks/light vehicles, air, modal split (passenger/goods).

MURE indicator scoreboard

Methodology

The MURE Policy Scoreboard presents different types of scoring, in particular scoring based on *outputs* (achieved energy savings). In that, MURE makes use of the unique feature that a large number of energy efficiency measures are quantified in the MURE database. By default the scoring period comprises measures from the year 2000 to present but other periods can be selected. The energy savings are those projected for 2020 at present. This is an important target year in the frame of the European Energy Efficiency Directive. This approach differentiates the MURE policy scoreboard from the ACEEE scoreboards and other approaches which are *input*-based (e.g. be comparing normalised subsidies amounts to energy efficiency or number of appliance standards).

Information on impacts in terms of energy savings for each measure in the MURE database takes two forms:

- Quantitative information from dedicated evaluations of measure impacts[5], mostly from evaluations at national level. Many large impact measures have individual evaluations
- Semi-quantitative expert estimates on measure impacts which group the measures in three categories. Measures saving less than 0.1 % of the sector energy consumption (low impact measures), measures saving 0.1 to less than 0.5 % of the sector energy consumption (medium impact measures), and measures saving more than 0.5 % (high impact measures). For measures in the cross-cutting database the percentages refer to the overall final energy consumption of the country. Nearly 90 % of all measures in the database have been classified in such a manner. These categories have been derived from the distribution of percentage savings in cases when a full quantification of measures is available. The expert estimates are done by national experts from the energy efficiency agencies participating in the project and are regularly verified and adapted, if necessary.

Ranking procedure

The following procedure is carried out for the ranking of the energy efficiency measures:

- Quantitative impacts are expressed as a percentage of the information sectoral final energy consumption. For measures in the cross-cutting database the percentages refer to the overall final energy consumption of the country (for the year 2010). It is assumed that measures with quantitative estimates already include the interaction with other measures.
- Semi-quantitative estimates are converted to quantitative estimates by using 0.1 %, 0.3 % (as the average of the category 0.1 to less than 0.5 %) and 0.5 % to characterize the measures. In order to consider interaction between those measures, a default interaction coefficient per measure is integrated into the calculation (derived from empirical information from the quantitative measures where interaction is explicitly considered).

5. This information is gathered in formal tables and can be retrieved for the policy scoreboard. At present around 40 % of all 2,400 policy measures in the MURE database have such a quantitative policy impact evaluation.

- Measures without a quantitative or semi-quantitative estimate are considered as "zero impact" in order to give a malus to measures which are not characterized. In such a manner, the monitoring practice in the country is also taken into account.
- Savings from the cross-cutting sector are established in a similar manner. The savings from the cross-cutting measures are then distributed over the four sectors (residential, services, transport and industry) according to the sector share in final energy consumption.

The sector results are then normalized with the sector shares in overall final energy consumption of the country which is set to 100 %. The corresponding graphs can then be read as percentage savings achieved with the energy efficiency measures of a country (compared to the energy consumption of the year 2010). For more details on the methodology see the ODYSSEE-MURE methodological documents.[6]

Data Collection and Requirements

The data on the energy efficiency policies (including on their quantitative impacts) are collected, similar to the indicator approach, in a regular manner by energy efficiency agencies in the EU28 Member States, Norway, Switzerland and Serbia. They collect in the database national evaluations of energy efficiency measures. Around 40 % of all measures in the MURE database have such quantified information.

DATA REQUIREMENTS AND METHODOLOGY OF THE ACEEE STATE AND INTERNATIONAL ENERGY EFFICIENCY SCORECARDS

International energy efficiency scorecard

Methodology

The 2016 edition of the *ACEEE International Energy Efficiency Scorecard* evaluates 23 of the world's top energy-consuming countries on 35 different metrics. The analysis examines efficiency in the three largest end-use energy categories: buildings, industry, and transportation. Also evaluated, as a separate category, are national efforts toward improving energy efficiency. Metrics used to evaluate countries in the *International Scorecard* are either policy or performance oriented. Policy metrics highlight best practices implemented by a country. They can be either qualitative or quantitative. Examples include national targets for energy efficiency, building and appliance labelling, and fuel economy standards for vehicles. The performance-oriented metrics measure the energy used per unit of activity or service extracted; they are quantitative. Examples include the efficiency of thermal power plants, energy intensities of buildings and industry, and average on-road vehicle fuel economy.

The point allocation for the 2016 edition is split 60/40 between policy and performance metrics based on feedback received on previous editions of the report and the understanding that performance metrics in part measure factors other than energy efficiency such as the local climate's impact on the degree to which buildings are heated or cooled. The metrics used to evaluate a country's progress and the number of associated points is provided in detail by Kallakuri et al. 2016 (p. 4/5).

The analysis in the 2016 International Scorecard is intended to provide readers with a broad bird's eye view comparison of energy use and the status of energy efficiency policies in each country. It is challenging to find a methodology that adequately captures energy efficiency efforts and allows for comparison across countries given inherent differences between countries. For instance, physical factors such as geographic size, climate, elevation, and availability of natural resources determine to a great extent the energy a country uses. These conditions are difficult to control for, and we were not always able to account for them in our scoring methodology. In general, we made only modest adjustments to raw data to enable basic comparisons across countries.

Ranking procedures

The maximum score a country can earn is 100. Each of the four metric categories is scored out of 25 points. The highest score available for a given metric is given to at least one country, which means that if any country were to emulate the top practices and results in every metric, it could obtain a score of 100. As a result, one of the primary goals of the International Energy Efficiency Scorecard is to provide countries interested in making progress on energy efficiency with a series of best practice cases that can be replicated.

Data collection and requirements

Whenever possible data and indicators on energy consumption and energy efficiency policy was collected from centralized, internationally recognized sources such as the IEA, the World Bank, the World Energy Council, the Organization for Economic Co-operation and Development (OECD), and the International Council on Clean Transportation (ICCT). This information was supplemented with country-level research by ACEEE staff. Collected data was reviewed by in-country and subject-matter experts to ensure accuracy and reliability by circulating data requests and during the external review of the report. The most significant limiting factor for the analysis was the availability of consistent, comprehensive data. In a few cases in which data were unavailable, we assigned scores based on our best estimates from related information and expert opinion. Metrics were also largely chosen to reflect data availability.

State Energy Efficiency Scorecard

Methodology

The *ACEEE State Energy Efficiency Scorecard* assesses state policies and programs that improve energy efficiency across six different policy areas: utility programs and policies, transportation, building energy codes and compliance, combined heat and power, state-led initiatives, and appliance and equipment standards. As with the International Energy Efficiency Scorecard, the State Scorecard also documents best practices, recognizes leadership and provides an annual benchmark of the progress of state energy efficiency policies.

To reflect the diversity in policy environments, demographics, and economic composition amongst states, we chose metrics that are flexible enough to capture the range of policy and program options that states use to encourage energy efficiency.

6. ODYSSEE-MURE methodological documents available here: http://www.indicators.odyssee-mure.eu/php/odyssee-scoreboard/documents/methodology-odyssee-scoreboard.pdf.

The policies and programs evaluated in the *State Scorecard* aim to reduce end-use energy consumption, set long-term commitments for energy efficiency, and establish mandatory performance codes and standards. They also help to accelerate the adoption of the most energy-efficient technologies, reduce market, regulatory, and information barriers to energy efficiency, and provide funding for efficiency programs. The metrics and associated point allocation are provided in ACEEE 2016.

With the exception of utility policies, the other policy areas are not scored on reported savings or spending data attributable to a particular policy action since these data are not widely available. Instead, given the lack of consistent ex post data, we have developed best-practice metrics for scoring the states. Although these metrics do not score outcomes directly, they credit states that are implementing policies likely to lead to more energy-efficient outcomes. For example, credit is given for *potential* energy savings from improved building energy codes and appliance efficiency standards since *actual* savings from these policies are rarely evaluated. Outcome metrics are reflected in the methodology where possible; for example, electric vehicle (EV) registrations and reductions in vehicle miles travelled (VMT) both represent positive outcomes of transportation policies.

Ranking procedures

We allocated points among the policy areas to reflect the relative magnitude of energy savings possible through the measures scored. We relied on an analysis of scholarly work and a variety of cross-sector potential studies in addition to expert feedback to inform our understanding of the energy savings available in each policy area and the point allocation. Of the 50 total points possible, 20 points were given to utility and public benefits program and policy metrics, 7 points to building energy codes, and 4 points to improved CHP policies. We used the same methodology to allocate the other policy area points, awarding 10 points for transportation policies and programs and 2 points for state appliance and equipment standards.

Data collection and requirements

As with the *International Scorecard,* data for each state is collected by ACEEE staff from a variety of centralized data sources and supplemented by state-specific research. This data is then verified and corroborated through a data request which gives state energy offices and utility commission officials the opportunity to correct and review data. Data collected is also displayed publicly on a corresponding website database.[7]

Comparison of scoreboard methodologies

There are major similarities but also considerable differences between the two main scoreboards considered in this article. Among the similarities between the ACEEE and ODYSSEE-MURE scoreboards one notes that both use a strong ranking principle, meaning both scoreboards weight the different criteria and add them up to an overall score. Both also allow transparently tracing the position in the ranking back to individual criteria at the level of individual energy uses or to policies. The degree of details depends on the availability of data. This may be different across countries, especially when it comes to international scoreboards.

There are also several differences between ACEEE and ODYSSEE-MURE approaches. While ACEEE uses a combined performance and policy ranking, ODYSSEE-MURE uses separate rankings for policy and performance. In the ongoing ODYSSEE-MURE project a further step will be taken towards unifying the performance scoreboard with the policy scoreboard. It will remain important, however, to present transparently the different components. Another difference is seen in the level of adjustments made to normalize raw data to enable comparison. The ODYSSEE-MURE approach uses a number of adjustments to compare the European countries. For example, the data are climate adjusted to take into account the climatic impacts. Further there are adjustments for differences in industrial structure, as for example a country with a lot of energy intensive industry (e.g. Germany or Finland), necessarily have a higher level of energy consumption than a country with a lot of service industries (e.g. Denmark). ACEEE also makes adjustments for climate using heating and cooling degree days in the buildings sector and accounts for energy intensity in the industrial sector, but generally avoids these adjustments unless absolutely necessary. A final difference is that ODYSSEY-MURE has a procedure for measuring the impact of particular policies on energy use to award points while ACEEE awards points only for the presence of best practice policies.

As evidenced by the suite of ACEEE and ODYSSEE-MURE scoreboards, there are numerous approaches that can be taken to evaluate energy efficiency at the national, state or local level. Above all they can provide useful results for the discussion, as long as the objective, target audience and methodology for a given scoreboard are defined in a transparent manner. Future scoreboard experiments will show which methodologies to harmonise and which show different facets of energy efficiency.

Results

RESULTS FROM THE ODYSSEE-MURE ENERGY EFFICIENCY SCOREBOARD

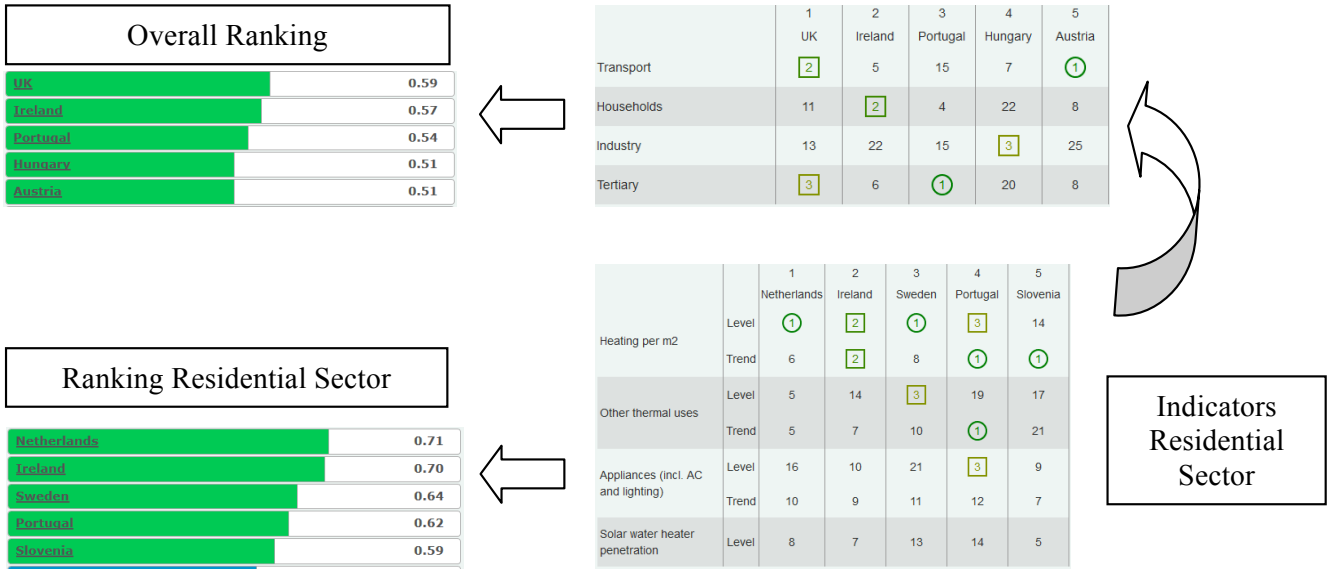
Results from the ODYSSEE scoreboard (performance)

Figure 1 shows most recent results from the ODYSSEE indicator scoreboard, combining levels and trends in a performance ranking, first by sector (here the household sector as an example) and finally the overall ranking. Important is, not only to provide an overall ranking but also the sub-rankings in the sectors and how they are composed, as this shows the weak points even for the leader of the performance ranking. The UK appears in the front position of the performance ranking; however, a look to the household sector shows that the UK is doing less well there (the data are adjusted for climate differences in Europe); there it does not appear among the top 5 but only at rank 11. Nevertheless, through good performance in the transport and service sector, it is overall performing better than other countries.

Results from the MURE scoreboard (energy efficiency policies)

The MURE ranking of energy efficiency policies may occur along the following four presentations of results:

7. See ACEEE State and Local Policy Database: http://database.aceee.org/.



	1 UK	2 Ireland	3 Portugal	4 Hungary	5 Austria
Transport	2	5	15	7	1
Households	11	2	4	22	8
Industry	13	22	15	3	25
Tertiary	3	6	1	20	8

		1 Netherlands	2 Ireland	3 Sweden	4 Portugal	5 Slovenia
Heating per m2	Level	1	2	1	3	14
	Trend	6	2	8	1	1
Other thermal uses	Level	5	14	3	19	17
	Trend	5	7	10	1	21
Appliances (incl. AC and lighting)	Level	16	10	21	3	9
	Trend	10	9	11	12	7
Solar water heater penetration	Level	8	7	13	14	5

Figure 1. Ranking energy efficiency trends and levels based on ODYSSEE indicators (performance ranking).

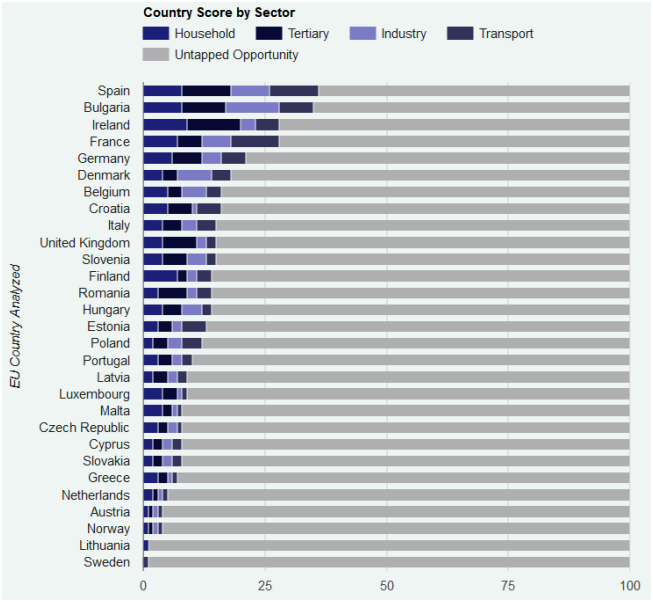


Figure 2. Ranking of energy efficiency polices based on ODYSSEE indicators (policy ranking).

- *Output-based scoring (based on energy savings):* This scoreboard makes use of the information in the MURE database on energy savings ("policy output") and compares the savings with the final energy consumption of the sector or total final energy consumption for a given year (at present 2010). This ranking is useful when it comes to the question, how much energy has been saved by measures taken since a certain date (here year 2000).
- *Output-based scoring (related to energy efficiency potentials):* This scoreboard compares the savings from measures since a recent date (here 2013, the starting year of the Energy Efficiency Directive EED) with the energy efficiency potentials at the time horizon 2020 or 2030. This allows to see which countries have most advanced in the realisation of their energy efficiency potentials with recent measures. The energy efficiency potentials were derived from a modelling study (Fraunhofer ISI 2014).
- Output-based scoring *(related to 2020 energy efficiency targets)*: This scoreboard compares the savings with the energy efficiency targets at the time horizon 2020. By default the scoring period comprises measures from 2013 to present. The targets are either calculated as a flat 20 % target for all EU countries (reflecting the overall 20 % target of the EU for energy efficiency) or by taking the EED targets provided by the different EU Member States.
- *Input-based scoring*: This input-based scoreboard makes use of the information in the MURE database on the inputs to energy efficiency policies (e.g. amount of final subsidies) and normalizes the inputs with respect to the size of the country (e.g. Gross Domestic Product or Population) if necessary or other relevant parameters. By default the scoring period comprises measures from 2000 to present.

Figure 2 shows results from the ranking of energy efficiency policies according to the savings achieved. The scale on the horizontal axis can be read as percentage savings achieved by policy measures since 2000 (a value of 25 for example implies that 2.5 % of final energy has been saved through energy efficiency policy measures. From a comparison with the performance indicators (which also include autonomous energy savings and savings from previous policies). The impacts can be traced by to the sectoral level, and then to individual policies.

RESULTS FROM THE ACEEE STATE AND INTERNATIONAL ENERGY EFFICIENCY SCORECARDS

International energy efficiency scorecard

Figure 3 shows results from the 2016 International Energy Efficiency Scorecard. Germany earned the most points, scoring 72.5 points, closely followed by Italy and Japan, tied in second place with 68.5 points. The lowest scoring country was Saudi

Arabia with 15 points, although this result was largely due to a lack of available data.

We see different leaders emerging when we rank countries according to policy-only metrics or performance-only metrics. Looking at policy-only metrics, Germany retains the #1 slot, but France joins Italy in the top three and Japan falls several rankings. For performance-only metrics, Japan rises to the top and takes the #1 slot, followed by the United Kingdom and Germany. Results clearly change depending on which metrics are applied and how they are weighted. A wide variety of approaches to metric selection, methodology, weighting can be used to produce valid results and useful rankings. For this reason, the factors that influence rankings must be clearly explained so the results can be interpreted.

State Energy Efficiency Scorecard

Figure 4 shows results from the 2016 State Energy Efficiency Scorecard. California and Massachusetts tied for the top spot in 2016 with Vermont and Rhode Island, followed by Connecticut and New York joining them in the top tier.

COMPARISON OF SCOREBOARD RESULTS

Focusing purely on energy use metrics comes with challenges. First, as mentioned above, energy use can be impacted by more than just energy efficiency. Geographic conditions and location, demographics, climate, economy, and other factors play a significant role in determining total energy consumption but are hard to control for when looking at energy intensity or energy use data. Comparing the percent change in energy use over time is often a good way to draw comparisons and reduce the need to account for some of these non-energy impacts. However, this approach can also introduce disadvantages. Results are highly sensitive to the baseline year selected and rankings become significantly influenced by the time period analyzed. For example, a jurisdiction that experienced an economy-wide recession during the time period chosen could rise in the energy-efficiency rankings without undertaking any activity to reduce energy use. Thus, incorporating an evaluation of policy efforts can often provide a good complement to energy use metrics.

Likewise, a scoreboard that looks simply at the presence of certain policies in given jurisdictions has its own limitations.

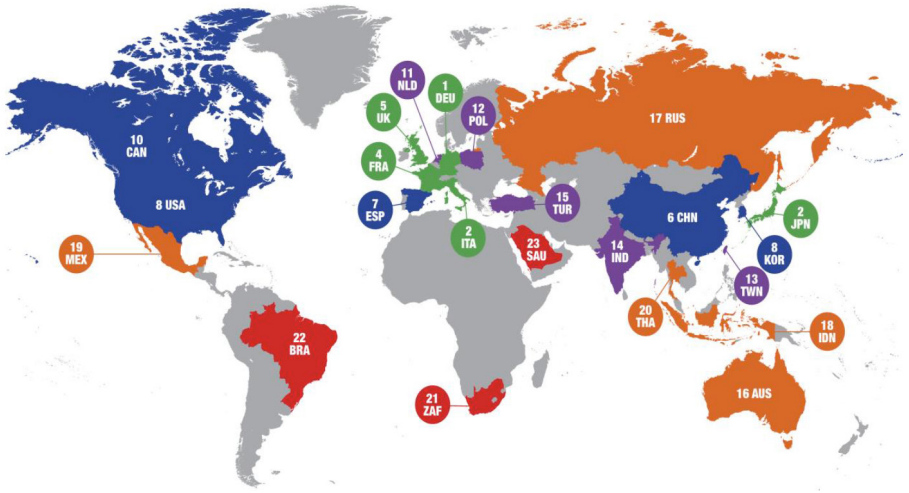


Figure 3. International ranking of energy efficiency performance and polices in the ACEEE International Energy Efficiency Scorecard.

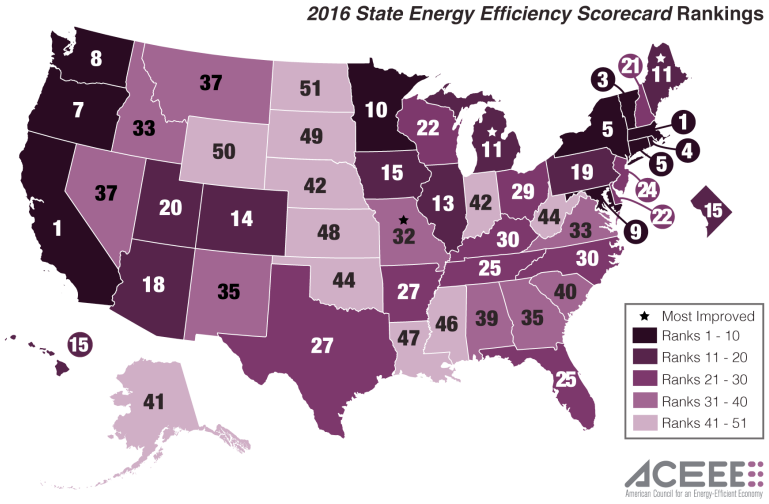


Figure 4. Ranking in the 2016 ACEEE State Energy Efficiency Scorecard.

While tracking particular energy efficiency policies can be a good indicator of how dedicated a country, state, or city is towards achieving energy reductions, policy-only scoreboards rarely give a good sense of the detailed differences in energy consumption by end use or where the best opportunities for energy efficiency lie. Additionally, the adoption of a policy often does not necessarily mean that it is fully implemented or adhered to. To fairly determine policy progress on energy efficiency, scoreboard methodologies should ideally evaluate implementation of each considered policy. However, this information is often hard to come by and usually requires on-the-ground research. This is a major step forward in the ODYSSEE-MURE project which gathers quantitative impacts from a large number of policies.

Another common observation is that room for improvement on energy efficiency exists everywhere. Countries do not earn maximum points in either ACEEE or ODYSSEE-MURE.

Future improvement and uses of energy efficiency scoreboards

The ability to make comparisons is an important exercise for assessing progress and identifying weaknesses in any field and presenting information in a ranking format can provide a useful framework for motivating future action. It is important to understand the way rankings are done and, as long as any limitations are recognized and discussed, they can prompt new ideas and changes in policies. To improve the use of energy efficiency scorecards in the future, the purpose of the ranking and the audience for the results should be clear. Efficiency scorecards are primarily tools for policy analysis and are not likely to be useful decision-making tools for other audiences like energy managers or efficiency engineers. A good ranking will be methodologically rigorous and rely on accurate data while also striking an even balance between the simplicity of identifying approaches to saving energy and the complexity of accurately measuring energy performance.

References

ACEEE, The 2016 State Energy Efficiency Scorecard, Research Report U1606, 26 September, 2016,http://aceee.org/research-report/u1606.

Bryant, Ph., Energy Works: Mississippi's Energy Roadmap, p.21, 2012, https://www.naseo.org/Data/Sites/1/documents/stateenergyplans/MS_20121.pdf.

DENEFF/Fraunhofer ISI, Energy efficiency progress: Germany only ranks in the middle in the EU despite high levels of energy efficiency and policy, 21 April 2015, http://www.isi.fraunhofer.de/isi-en/service/presseinfos/2015/press-release-12-2015-energy-efficiency-progress-policy.php.

Press echoes see for example (in German): http://www.elektrotechnik.vogel.de/andere-eu-staaten-ueberholen-deutschland-in-der-effizienz-a-486541/; http://www.energynet.de/2015/04/22/energieeffizienz-deutschland/; http://www.hannovermesse.de/de/news/energieeffizienz-ist-die-zweite-saeule-der-energiewende.xhtml.

Eldridge, M., Prindle, B., York, D. and Nadel , S., The State Energy Efficiency Scorecard for 2006, Report Nb. E075, June 2007, http://aceee.org/sites/default/files/publications/researchreports/e075.pdf.

European Commission, Commission Staff Working Document Impact Assessment Accompanying the Document Proposal for a Directive of the European Parliament and of the Council amending Directive 2012/27/EU on Energy Efficiency, Part 2/3, SWD(2016) 405 final, Brussels, 30 November 2016, http://ec.europa.eu/energy/sites/ener/files/documents/1_en_impact_assessment_part2_v5.pdf.

Kallakuri, Ch., Vaidyanathan, S., Kelly, M. and Cluett, R., The 2016 International Energy Efficiency Scorecard, Report E1602, American Council for an Energy-Efficient Economy, July 2016.

Learning is Open, http://learningisopen.org/oecd/germany.html.

ODYSSEE-MURE, Indicators Scoreboard Facility, http://www.indicators.odyssee-mure.eu/energy-efficiency-scoreboard.html. Methodological document at http://www.indicators.odyssee-mure.eu/php/odyssee-scoreboard/documents/methodology-odyssee-scoreboard.pdf. Policy Scoreboard, http://www.measures-odyssee-mure.eu/scoreboard-energy-efficiency-policy.asp. Methodological document at http://www.measures-odyssee-mure.eu/MureScoreboardMethodology.pdf.

OECD/JRC, Handbook on Constructing Composite Indicators, Methodology and User Guide, Organisation for Economic Co-Operation and Development OECD, Paris, 2008. http://www.oecd.org/std/leading-indicators/42495745.pdf.

Office of Commissioner Brandon Presley, News Release: Presley Joins National Leaders to Announce Mississippi as Most Improved Energy Efficiency State in Nation, 6 November, 2013. http://www.psc.state.ms.us/commissioners/northern/press%20releases/2013/Presley Announces Mississippi as Most Improved Energy Efficiency State in the Nation.pdf.

Pennsylvania, 2015 Climate Change Action Plan Update, Department of Environmental Protection, Pennsylvania, August 2016, http://www.elibrary.dep.state.pa.us/dsweb/Get/Document-114163/FINAL 2015 Climate Change Action Plan Update.pdf.

PLAN, Transforming Los Angeles, p. 102, 2015. https://www.lamayor.org/sites/g/files/wph446/f/landing_pages/files/The%20pLAn.pdf.

State of New Hampshire, Public Utilities Commission, DE 15-137, Gas And Electric Utilities, Energy Efficiency Resource Standard, Order Approving Settlement Agreement, Order No. 25,932, 2 August, 2016, http://www.puc.state.nh.us/Regulatory/Docketbk/2015/15-137/ORDERS/15-137_2016-08-02_ORDER_25932.PDF.

Acknowledgements

The ODYSSEE-MURE project acknowledges the support from the European Commission (EASME) under the Intelligent Energy for Europe Programme and the H2020 programme.