



ERNEST ORLANDO LAWRENCE BERKELEY NATIONAL LABORATORY

International Database of Efficient Appliances (IDEA): A Novel Tool for Efficiency Program Development and Evaluation

Brian F. Gerke, Ph.D., Michael A. McNeil, Ph.D., and
Thomas Tu

Lawrence Berkeley National Laboratory
Berkeley, CA 94720

June 2015

This work was supported by the Laboratory Directed Research and
Development Program of Lawrence Berkeley National Laboratory
under U.S. Department of Energy Contract No. DE-AC02-05CH11231.

DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor The Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or The Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or The Regents of the University of California.

COPYRIGHT NOTICE

This manuscript has been authored by an author at Lawrence Berkeley National Laboratory under Contract No. DE-AC02-05CH11231 with the U.S. Department of Energy. The U.S. Government retains, and the publisher, by accepting the article for publication, acknowledges, that the U.S. Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for U.S. Government purposes.

International Database of Efficient Appliances (IDEA): A Novel Tool for Efficiency Program Development and Evaluation

*Brian F. Gerke, Ph.D., Michael A. McNeil, Ph.D., and Thomas Tu,
Lawrence Berkeley National Laboratory*

ABSTRACT

Energy efficiency standards and labeling (S&L) programs for appliances are an important part of most developed economies' energy policy portfolios. It is difficult to evaluate the full impact of these programs on global energy demand, however, since such programs may have significant additional effects beyond the borders of the market being regulated. At present, the data required to understand the international impacts of S&L programs are highly fragmented, with different required data residing with a wide variety of retailers, manufacturers, market research firms and government agencies. Further complicating matters, efficiency data is measured using different metrics and test procedures from nation to nation. Until the various data sets can be brought together and made comparable, a full accounting of S&L impacts will remain elusive. To address these problems, Lawrence Berkeley National Laboratory is constructing the International Database of Efficient Appliances (IDEA), a flexible database for organizing, aggregating and storing market and energy-performance data on a variety of appliances, electronics, and other equipment sold worldwide, collected via web-crawling and other data collection techniques. This paper describes the concept behind IDEA, the technical and analytical aspects associated with its development, and example applications in the context of efficiency program evaluation. A first look at early IDEA data on refrigerators in India is presented.

Introduction

Most large, industrialized nations have policies to promote energy efficiency in appliances and other consumer goods, with the goal of reducing energy consumption and/or greenhouse gas (GHG) emissions. Minimum energy performance standards (MEPS) programs, such as the Appliance Standards* program at the US Department of Energy (DOE), set minimum efficiency levels that various products must meet to be legally sold. Labeling programs, such as FTC's mandatory Energy Guide† or DOE/EPA's voluntary ENERGY STAR‡, provide retail consumers with information about appliance energy efficiency. Standards and labeling (S&L) programs are considered an important part of the global response to climate change; in particular, more stringent standards are a central component of the President's climate action plan (White House 2013).

It is difficult, however, to evaluate the full impact of S&L programs on global energy demand and GHG emissions. The development of S&L programs is typically based on estimates of the energy savings and avoided emissions that will occur within the market being regulated, but such programs may have significant effects on international markets as well. For example, it is plausible to imagine that S&L programs drive innovations that lower the cost of efficiency worldwide. Alternatively, they may eliminate inefficient products in regulated

* <http://energy.gov/eere/buildings/appliance-and-equipment-standards-program>

† <http://www.consumer.ftc.gov/articles/0072-shopping-home-appliances-use-energyguide-label>

‡ <http://www.energystar.gov>

markets but drive them toward less regulated markets in the developing world, thus boosting efficiency in some countries while reducing it in others. Depending on the relative size of such different cross-market effects, the global impacts of S&L programs could be much larger or much smaller than current estimates indicate. Moreover, if S&L programs affect the availability of efficient products in the developing world, this will have implications for the sustainability of global development.

At present, the data required to understand the global impacts of S&L programs are almost hopelessly fragmented and disjoint. Data on product price, availability, and market share reside with a wide variety of retailers, manufacturers, and market research firms. Energy efficiency data is located in multiple government databases and is measured using a variety of metrics and test procedures, complicating the comparison of efficiency across markets. Until these various data sets can be brought together and made comparable, a full accounting of S&L impacts will remain elusive.

To improve this situation and to lower the barriers to data-driven evaluation of appliance S&L programs, Lawrence Berkeley National Laboratory (LBNL) is constructing the International Database of Efficient Appliances (IDEA), a flexible database for collecting, cleaning, aggregating, and storing data on a wide variety of energy-consuming appliances from markets across the globe. IDEA aggregates, correlates, and regularizes data from online retailers, manufacturer websites, and government databases, to yield a composite global record for each item in the database. Using web-crawling techniques, data can be collected at regular intervals to facilitate the real-time tracking of market shifts that may arise from changes in efficiency policies and technologies. With its global scope and high-frequency data collection, IDEA enables novel approaches to rapid evaluation for S&L and other efficiency programs.

This paper proceeds as follows. First, we lay out the technical goals and specifications of IDEA, including the data fields needed for meaningful analysis, potential data sources, strategies for aggregating them, and challenges arising from data gaps and incompatibilities. Next, we outline some of the potential applications for IDEA. These include ex-post evaluation of program impacts, large-scale market tracking, and more sophisticated quantitative analysis of market effects such as pricing policies, technological learning and policy-driven innovation. We then briefly summarize the IDEA framework and basic functionality and provide some first-look results on refrigerators in India, using data from a prototype of IDEA. Finally, we briefly discuss the scalability of this type of data to a broader range of markets and energy end-uses.

Goals and applications for an international appliance database

The vision for IDEA is of a comprehensive database of energy consumption and price information for appliances mass-marketed on the Internet, across a broad range of international markets. The database contains detailed information on the individual-model level, including

- Identifying information (e.g., model numbers, brands, and unique product identifiers such as UPC or EAN codes)
- Current and historical price data
- Product features and attributes
- Data on the availability and popularity of products at various efficiency levels, within various markets and over time
- Energy efficiency data certified by governments or other reputable sources.

- Region-specific data, such as currency, units of measure, and particular energy efficiency metrics and test procedures.

When fully implemented, IDEA has the potential to serve as a global resource for data on appliance efficiency and for research into the interactions between appliance efficiency, price, market penetration, and energy policy measures. To construct IDEA, it is necessary to draw together data from a variety of disparate sources, including online retailers, manufacturer websites, and efficiency databases maintained by governments and NGOs. IDEA leverages these existing sources of public-facing data, presently dispersed across the Internet, and combines them into the rich, comprehensive dataset envisioned above. Such a dataset would be valuable in a wide variety of contexts in energy program evaluation and beyond. Potential applications of IDEA for S&L program evaluation include the following:

- *Tracking of availability and efficiency trends.* Appliance S&L or incentive programs aim to increase the typical efficiency of appliances sold in a particular market, either by banning inefficient models (in the case of MEPS), or by incentivizing the purchase of efficient ones. A central component of evaluating the impact of such programs is the tracking of products offered for sale, to gauge the speed at which banned products exit the market, and the extent to which incentivized products increase their market presence. IDEA provides a ready-made tool for tracking such market shifts in real time. It also enables the computation of aggregate indicators of characteristic efficiency within a market
- *Tracking of price trends.* Some efficiency programs, such as MEPS, are expected to result in increased typical prices for products, which are supposed to be fully offset by energy-cost savings. By using IDEA to track appliance prices before and after the introduction of a new policy, it will be possible to determine the actual change in price that occurs at the commencement of the new policy. Interestingly, recent retrospective research has suggested that some efficiency S&L programs may unexpectedly have yielded price declines, rather than increases (Spurlock 2013; Van Buskirk, Kantner et al. 2014). A tool for detailed price tracking tool like IDEA would allow such possible effects to be investigated in more detail at the time of program implementation. For an example of using web-crawling data to track the price of efficient products, see a recent study of LED light bulb prices by Gerke et al. (2014).
- *Tracking of online popularity.* Many Internet retail sites present data that can be used to gauge the relative popularity of different appliance models, such as the rank of each model's sales among all products sold on the site, or the total number of user reviews. Recent research suggests that such data can be converted into a reasonably accurate proxy for market share (Touzani & Van Buskirk 2015). Such data could then be used to track shifts in relative popularity, as a function of efficiency, in response to labeling programs.
- *Market surveillance for compliance verification.* A crucial evaluative component of any appliance S&L program is the assessment of program compliance by manufacturers and retailers. Typically, manufacturers are required to certify their products' energy consumption to the authority managing a labeling program (for example) before applying an efficiency label. A tool like IDEA, which cross-references retail models against the resulting certification database, provides real-time support for compliance

- evaluation by confirming that models displaying the efficiency label are consistent with the certification database.
- *International efficiency benchmarking.* The international scope of IDEA will allow the development of international efficiency benchmarks that are applicable to a broad range of global markets. Benchmarking provides a key indicator to program managers about the efficiency of their own markets relative to international norms, as well as the relative stringency of their regulations. Historically, this data has been difficult to access, although recently great progress has been made in this area by the Collaborative Labeling and Appliance Standards Program (CLASP) and by the International Energy Agency[§].
 - *Estimation of inter-market program impacts.* It is difficult to evaluate the full impact of S&L programs on global energy demand and GHG emissions, since evaluation is typically focused on the energy savings and avoided emissions that will occur within the market being regulated. However, such programs may have significant indirect effects on international markets as well. For example, a recent study showed that efficiency programs in the EU have had significant impacts beyond the EU member states through emulation of EU policies by other nations (Waide et al. 2014). The inter-market impact of developed nations' programs may go beyond policy emulation, however, since S&L programs in large markets can affect the mix of products available in smaller importing markets. These impacts may be positive, if they eliminate inefficient products or increase the availability of efficient products in smaller markets; or they may be negative, if inefficient products banned in large markets are dumped in smaller, unregulated markets. IDEA will allow the first systematic accounting of these indirect international effects of national S&L programs.

IDEA also has many potential applications that go beyond S&L program evaluation. First and foremost, it will serve as a powerful tool for efficiency program development by facilitating the determination of baseline efficiency distributions and potential efficiency improvements. The price-tracking functionality will also help to support the incorporation of price trends into efficiency-policy analyses, as discussed by Desroches et al. (2012). Additionally, IDEA may facilitate novel approaches to program implementation. For example, small importer nations could use IDEA to determine whether incoming appliances meet the MEPS of relevant exporting nations. This allows an extremely low-cost approach to MEPS for small and developing nations, which leverages the programs of larger nations, and which can be enforced simply at the point of import via a bar-code scanner linked to IDEA. The database can also support consumer-facing applications that could increase the reach of efficiency labeling programs by allowing consumers to access labeling information by scanning a bar code or QR code. This could be particularly effective in developing nations that position government efficiency labels as a sign of product quality, *e.g.*, as in Ghana^{**}.^{††} Finally, IDEA could be used to support the identification and sourcing of ultra-efficient appliances for use in off-grid or micro-grid applications in the developing world.

[§] See <http://clasponline.org> and <http://mappingandbenchmarking.iea-4e.org>

^{**} <http://energycom.gov.gh/Renewable/ghana-appliance-energy-efficiency-standards-and-labelling-programme.html>

^{††} <https://www.youtube.com/watch?v=DKRJdfp0c5A>

Strategies and challenges

The strategy for collecting the underlying data for IDEA is as follows. Model-level data on identity, price, attributes, etc., are collected directly from online retail or manufacturer websites using various automated data-collection techniques, as discussed below. We will refer to such data generally hereafter as *retail data*. Energy efficiency information is usually imported from public databases maintained by government agencies or by NGOs in association with mandatory or voluntary energy efficiency policies or programs. Examples of government databases include the ENERGY STAR database^{‡‡} hosted by the US Environmental Protection Agency, or the energy efficiency star-rating database hosted by India's Bureau of Energy Efficiency (BEE)^{§§}. An example of an NGO database is the product database maintained by the Consortium for Energy Efficiency for its efficiency-tier program^{***}. The data in the government and NGO databases is particularly valuable, since it is measured according to a well-defined test procedure and certified by the authority managing the database. We will refer to such data generally hereafter as *certification data*.

Where possible, data is imported directly from public-facing *application programmer interfaces* (APIs), such as the Best Buy Developer Portal in the US^{†††}, which allow programmatic access to the data underlying individual product listings on a web site. When APIs are not available, or are not sufficiently comprehensive, data can be collected using specialized web crawling programs that systematically visit the product listings on a retail or manufacturer site and parse them to extract the details for each product. The retail data are then aggregated on the individual-model level and cross-referenced against the certification data, yielding a composite record for each individual model in the database. By collecting data from Internet sources on a regular cadence, and cross-referencing them against certification data, IDEA will enable both real-time tracking and retrospective analyses of individual product prices and availability, as well as the development of broad, market-level efficiency indicators.

There are, however, numerous challenges that must be overcome before such a database can become a reality. We describe the major barriers below, and in the next section we discuss the ways in which IDEA overcomes each of them.

- *Product-collection flexibility.* The IDEA vision is of a database that contains data on a wide variety of disparate appliance types: refrigerators, air conditioners, ceiling fans, dishwashers, clothes washers, water heaters, furnaces, televisions, computers, and so on. Each product has different attributes relevant to understanding its energy profile. Internal volume is important for refrigerators and clothes washers but is irrelevant for televisions or ceiling fans, for example. Screen size is crucial for televisions, but meaningless for water heaters. IDEA must be sufficiently flexible to allow collection and storage of arbitrary data fields for each separate category of products stored in the database.
- *Non-standardized naming conventions.* Different data sources may use different terminology to refer to the same attribute of a given product. For example, terms such as *volume*, *capacity*, *storage space*, or *size* may all be used to refer to the same attribute

^{‡‡} <http://www.energystar.gov/products/certified-products>

^{§§} <http://beestarlabel.com/Home/Searchcompare>

^{***} <http://www.cee1.org/content/cee-program-resources>

^{†††} <https://developer.bestbuy.com/>

of a refrigerator. To effectively combine data from these sources, the database must recognize these different naming conventions and map them to a single data field with a standardized name.

- *Varying data field requirements.* As mentioned above, each product category has a different set of features that are of interest in IDEA. Moreover, this set of features may change over time, as new analytical techniques are applied to the data, or as entirely new features enter the market. For example, today it may be of interest to know whether a television has a curved screen, but this feature was unheard of only a few years ago. Therefore, IDEA must be flexible enough to allow arbitrary expansion of the set of features it collects and stores.
- *Region definitions.* A database that combines retail and certification data with an international scope raises definitional questions about the economic regions in which products are considered to be available. For example, the ENERGY STAR certification database contains data on models sold in both the US and Canada, but retail products are generally offered for sale on a country-by-country basis—e.g. at amazon.com and amazon.ca—to account for import restrictions or differences in language or currency. It is necessary in this case to consider appliances sold in North America both together (for comparison to the ENERGY STAR database) and separately (for determination of which models are available on the US and Canadian markets, respectively); the database must be sufficiently general and flexible in its region definitions to handle this situation.
- *Intra-region model duplication.* For each product type, it is likely that popular models will be listed on multiple different retail and manufacturer sites within each region. These duplicate listings must be recognized and combined in the final database to yield a unified record for each model.
- *Inter-region model duplication.* It is possible for the same model to be sold in multiple different markets; however identifying information such as brand name and model number may vary from region to region for otherwise identical models. Moreover, the values of various product attributes may vary from region to region, owing to differences in units of measure or specified measurement procedures within each market. The database must have the capacity to handle such differences.
- *Generalized certification records.* It is common for multiple models from a manufacturer's product line to have identical energy-consumption profiles (for example if they differ only in attributes that do not impact energy usage, such as color). Certification databases sometimes allow manufacturers to report model numbers in a generalized form, such that each reported record represents multiple products, all having the same energy efficiency. For example, a manufacturer might report certain characters in their model numbers as asterisks to indicate that these positions in the model number can take various values. The algorithms used for generalizing model numbers vary from database to database and from manufacturer and manufacturer. A database that aims to cross correlate model-level retail data against generalized certification databases must have a robust and flexible methodology for reverse-engineering the model-number generalization algorithms and applying them to the retail model numbers.

IDEA framework and functionality

IDEA is designed to overcome all of the challenges laid out in the previous section. It is based on a flexible data framework that was constructed as part of the Super-Efficient Appliance Deployment^{***} (SEAD) initiative's data access project (Katzman, McNeil & Gerke 2013; Katzman, McNeil & Pantano 2014). The framework was specifically designed with the aim of enabling a database that was flexible and general enough to store data on efficiency and features for a wide variety of different appliance types across multiple different markets. It specified a solution for correlating and combining data from intra-regional duplicate model listings. It also proposed a process for specifying and extending canonical naming conventions for product features, as well as for conforming individual retailers' naming conventions to these canonical names. Thus, the SEAD framework provides solutions to the problems of product-collection flexibility, varying data field requirements, non-standardized naming conventions, and intra-region duplication described above.

IDEA is the first fully-featured, large-scale implementation of the SEAD data access framework. It also includes various extensions to that framework to handle the problems of region definition, inter-region model duplication, and generalized certification data. In the following, we briefly describe the steps involved in collecting and storing data in IDEA and note how this process yields a data set that overcomes each of the challenges described in the previous section. The components of the IDEA data flow are as follows:

- *Data feeds and economic regions.* All data stored in IDEA enters the database through a conduit known as a *feed*, which can be a retail website's API, a piece of web-crawling software collecting data from a retailer's or manufacturer's website, or an external product certification database maintained by a government or NGO. Each feed yields data on products that are available within a particular *economic region*, which comprises one or more countries. To allow for feeds that cover regions of varying scope, each country can be assigned to more than one overlapping economic region. Thus, for example, a feed for www.amazon.ca would be associated with an economic region that includes only Canada, and www.amazon.com would be associated with an economic region containing only the US, while the ENERGY STAR database feed would be associated with an economic region that includes both the US and Canada. In the normalization and cross-referencing steps (described below), data from each feed is correlated with data from all economic regions having a non-empty intersection (in country space) with the feed's region. In that case, data from amazon.com and amazon.ca would both be correlated against the ENERGY STAR database but not against each other.
- *Raw data collection and archiving.* All data enters the database from its feed in one of the common standard formats for data transmission over the web—either JSON or XML, depending on the feed. The raw data are then archived permanently in the database before further processing, so that additional data of interest may be extracted at will in the future.
- *Attribute mapping.* For each feed, data fields of interest are identified in the raw data, and their field names are mapped to the corresponding canonical field names defined within IDEA. This allows data from each feed to be assigned appropriately to the

^{***} <http://www.cleanenergyministerial.org/Our-Work/Initiatives/Appliances>

- canonical IDEA fields. Additional mappings can be added as needed to allow extraction of additional data fields.
- *Product normalization.* To aggregate duplicate listings of individual models within an economic region, all data is passed through a normalization step, which identifies such duplicates and combines their data into a single, unified record for each product. When a particular model already exists in the database, its information is updated with the incoming data, as appropriate, and the incoming price information is added to the product's price history.
 - *Attribute extraction.* Once duplicates have been combined, the attribute mappings defined for each feed are used to extract feature information from each of the feeds to populate the product record. To handle potential data conflicts (e.g., different web sites' reporting inconsistent product features), each feed is assigned a priority that is used to select which data to save in the database.
 - *Certification cross-referencing.* Because individual records in the certification databases may refer to multiple models, certification data is not passed through the product normalization step. Instead, the normalized retail data are cross-referenced by model number against the certification data, using sophisticated matching algorithms to account for the generalized model number information present in many certification databases.
 - *Inter-regional product matching.* The resulting set of processed data records includes one record for each unique model per economic region. It is possible that the same model may be sold in multiple economic regions. As discussed above, regional variations in language and data standards lead us to keep one record per region for each model; however, IDEA does allow identical products to be linked across regions, so that users can explore data from different regions for each model, where available.
 - *Data export for analysis.* Once the above steps are completed, analysis-ready data can be exported in a tabular form, with one row for each unique model (per region) and one column for each product feature that has been defined and mapped in the attribute mapping process. Price histories for individual products can also be exported.

Prototype database and example results

At present, we have developed a prototype of IDEA, with scope limited to a small number of product categories from a few markets. **Figure 1** shows an example of the data display for a small subset of refrigerator data from India and Mexico. Selling prices are shown both in the local currency and converted to US dollars.

It is interesting to consider the information this simple tool can deliver in the context of market surveillance for verifying compliance within a labeling program. As a simple example, we considered data for refrigerators sold online in India. After collecting the retail data, we cross-referenced it against the BEE star-rating database^{§§§} which yielded a BEE-certified star rating for each model that appears in the BEE database (labeling is only mandatory for frost-free refrigerators and is voluntary for other types, so not all models are expected to be present). Where available, we also extracted the star ratings that were advertised on the retail website for each model.

^{§§§} <http://beestarlabel.com/Home/Searchcompare>

141 total. Show: 10 25 100 all

<input type="checkbox"/>	Id	Brand	Mpn	Name	Url	Updated	Latest Sale Price	Latest Sale Price Currency	Latest Sale Price Converted
<input type="checkbox"/>	2216	Samsung	WA18F7L8DTA	Refrigerador 14 Pies Cubicos Samsung Silver RT38FEAKDSL/PREMIUM	Link	03/27/2015 11:17 a.m.	—	—	—
<input type="checkbox"/>	2129	GE	TA04Y07SS1	Frigobar 4 Pies Cubicos General Electric Plata	Link	03/27/2015 11:17 a.m.	4159.00	MXN	272.75
<input type="checkbox"/>	934	Samsung	RT54H667ESL	Samsung RT54H667ESL 528 L Double Door Refrigerator Clean Steel, Silver	Link	03/27/2015 11:17 a.m.	57040.00	INR	912.52
<input type="checkbox"/>	1210	Samsung	RT54H6679SL/TL	Samsung RT54H6679SL/TL Frost-free Double-door Refrigerator (528 Ltrs, 3 Star Rating, ...	Link	03/27/2015 11:16 a.m.	65500.00	INR	1047.86
<input type="checkbox"/>	2184	Samsung	RT53H6631SL/EM	Refrigerador Top Mount Samsung 19 Pies Cubicos Clean Steel Mod. ...	Link	03/27/2015 11:17 a.m.	14390.00	MXN	943.70
<input type="checkbox"/>	2153	Samsung	RT53H6631SL	Refrigerador 19 Pies Cubicos Samsung Acero Inoxidable RT53H6631SL	Link	03/27/2015 11:17 a.m.	13599.00	MXN	891.83
<input type="checkbox"/>	1026	Samsung	RT50H5809SL/TL	Samsung RT50H5809SL/TL 496 L Double Door Refrigerator Real Stainless Steel	Link	03/27/2015 11:18 a.m.	59990.00	INR	959.71
<input type="checkbox"/>	1211	Samsung	RT50H5809SL	Samsung RT50H5809SL Frost-free Double-door Refrigerator (496 Ltrs, 3 Star Rating, ...	Link	03/27/2015 11:16 a.m.	60000.00	INR	959.87
<input type="checkbox"/>	948	Samsung	RT49H567ESL	Samsung RT49H567ESL 481 L Double Door Refrigerator Clean Steel, Silver	Link	03/27/2015 11:17 a.m.	51622.00	INR	825.84

Figure 1. Example data from the prototype of IDEA. Shown are data for refrigerators sold in India and Mexico. Selling prices are shown in Indian Rupees (INR) and Mexican Pesos (MXN), as appropriate; and they are also shown converted to US Dollars, for comparison.

Figure 2 shows a comparison of the certified and advertised star ratings for the retail models that were found in the BEE database. Each panel displays a histogram of the models that are certified at a given star level in the BEE database. The histograms show the distribution of star values that were advertised online for these models. (No five-star certified models were found with advertised star ratings.) Notably, a majority of the models certified at the low-efficiency two-star level are being advertised online as higher-efficiency three, four, or five-star appliances. In addition to this apparent star-rating inflation, we found that 51% of the fridges we collected online in India were not present in the BEE certification database at all. As noted previously, efficiency labeling is not mandatory for all refrigerators in India, so some non-zero fraction of uncertified models is expected. However, of all the models advertised at the maximum five-star efficiency level, none were found to have five-star counterparts in the BEE database, rendering these high-efficiency claims suspect at best. The apparent inflation in the efficiency star ratings does suggest that at least some manufacturers or retailers believe that a high star rating is an attractive feature to refrigerator consumers in India, demonstrating a success of the BEE labeling program; however, this success is also apparently being undermined to some degree by star-rating inflation.

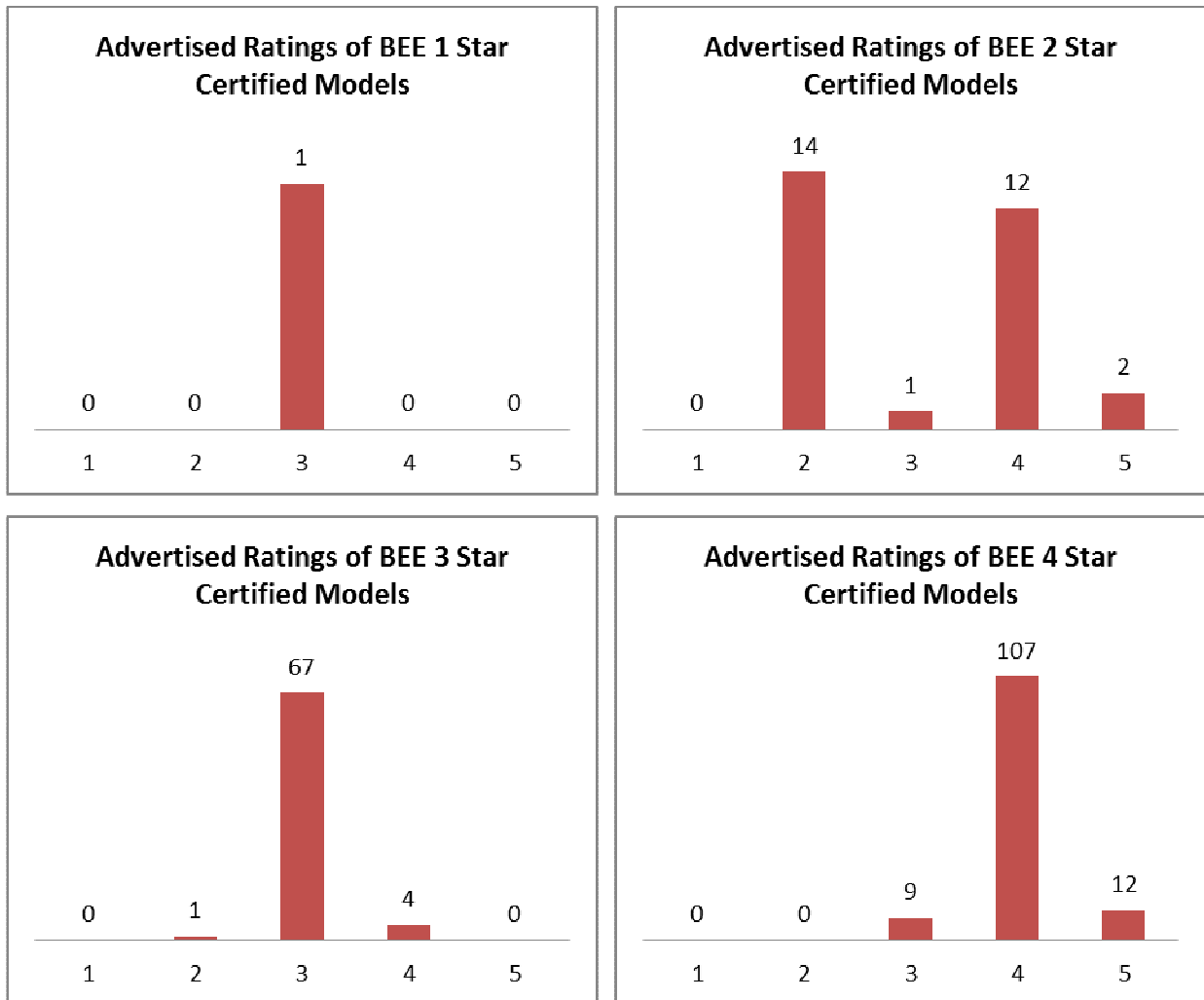


Figure 2. A comparison of the certified and advertised efficiency star ratings for refrigerators sold online in India. Each panel shows a histogram of models that are certified by BEE at a particular star level below the maximum five stars. (No five-star certified models were found with advertised star ratings.) The histograms show the star levels that were advertised online for fridges at each level. Notably, a majority of the models certified at the two-star level are advertised at a higher level.

Conclusions and future work

We have described the concept and potential applications of IDEA, which is a flexible database for collecting, correlating, managing, and storing data on the price, features, and efficiency of appliances gathered from various disparate sources across a wide range of international markets. There are numerous challenges to constructing such a database, all of which IDEA has been specifically designed to overcome. A working prototype of IDEA has been implemented and is now collecting data with limited scope.

We have used data on Indian refrigerators, collected using the prototype, to demonstrate the usefulness of IDEA in support of S&L program evaluation. Specifically, we explored the advertising of India's energy-efficiency star ratings as compared to the values certified to the

Indian government. We find a significant discrepancy for models with low ratings: low-efficiency refrigerators appear to be frequently marketed with an inflated efficiency rating.

The next steps are to expand IDEA data collection to a wider range of regions and appliances. This will enable a broad array of applications in international energy policy development, analysis, and evaluation. In the longer term, the vision is for IDEA to evolve into *IDEE*: The International Database of Equipment Efficiency. The database framework is sufficiently flexible that it could also encompass vehicles, building materials, solar panels, micro-inverters, and a wide variety of other equipment, in addition to appliances. IDEA/*IDEE* has the potential to serve as a global clearinghouse for efficiency data, utilized by governments and energy researchers worldwide.

References

- Desroches, Louis-Benoit, Karina Garbesi, Colleen Kantner, Robert D. Van Buskirk, and Hung-Chia Yang. 2013. "Incorporating Experience Curves in Appliance Standards Analysis." *Energy Policy* 52: 402-416.
- Gerke, B. F., Ngo, A. T., Alstone, A. L., and Fisseha, K. 2014. "The Evolving Price of Household LED Lamps: Recent Trends and Historical Comparisons for the US Market." LBNL Report No. 6854E. <http://eetd.lbl.gov/publications/the-evolving-price-of-household-led-l>. Berkeley: Lawrence Berkeley National Laboratory.
- Katzman, A., McNeil, M., and Gerke, B. "SEAD Energy Efficiency Data Access Project: Final Report." 2013. <http://www.superefficient.org/~media/Files/SL%20Project%20Reports/SEAD%20Data%20Access%20Report/SEAD%20Data%20Access%20Final%20Report.pdf>. Washington, DC: SEAD.
- Katzman, A., McNeil, M., and Pantano, S. 2014. "The benefits of creating a cross-country data framework for energy efficiency." *Proceedings of the 7th International Conference EEDAL'2013 - Energy Efficiency in Domestic Appliances and Lighting*. Luxembourg: Publications Office of the European Union.
- Spurlock, C. A. 2013. "Appliance Efficiency Standards and Price Discrimination." LBNL Report No. 6283E. <http://eetd.lbl.gov/publications/appliance-efficiency-standards-and-pr>. Berkeley: Lawrence Berkeley National Laboratory.
- Touzani, S. and Van Buskirk, R. 2015. "Estimating Sales and Sales Market Share from Sales Rank Data for Consumer Appliances (Draft)." LBNL Report No. 6989E. <http://eetd.lbl.gov/publications/estimating-sales-and-sales-market-s-0>. Berkeley: Lawrence Berkeley National Laboratory.
- Van Buskirk, R., Kantner, C. L. S., Gerke, B. F., and Chu, S. 2014. . "A retrospective investigation of energy efficiency standards: policies may have accelerated long term declines in appliance costs." *Environmental Research Letters* 9 (11): 114010.
- Waide, P., Janeiro, L., Surmeli, N., Gardiner, A., Tait, J., Fonsesca, P., Fong, J., Quaresma, N., and Evans, C. 2014. "Impact of the EU's Ecodesign and Energy/Tyre labelling legislation on third jurisdictions." https://ec.europa.eu/energy/sites/ener/files/documents/201404_ieel_third_jurisdictions.pdf. Ecofys: Utrecht.
- White House. 2013. "The President's Climate Action Plan." <http://www.whitehouse.gov/sites/default/files/image/president27sclimateactionplan.pdf>. Washington, DC: Executive Office of the President.