Applied Energy 205 (2017) 453–464

Contents lists available at ScienceDirect

Applied Energy

journal homepage: www.elsevier.com/locate/apenergy

# The International Database of Efficient Appliances (IDEA): A new tool to support appliance energy-efficiency deployment

Brian F. Gerke*, Michael A. McNeil, Thomas Tu[1]

*Lawrence Berkeley National Laboratory, 1 Cyclotron Rd. MS 90R4000, Berkeley, CA 94720, USA*

## HIGHLIGHTS

- IDEA collects, organizes, and stores data on appliance features and efficiency.
- Data can be gathered for any appliance type on any international market.
- IDEA enables new approaches to efficiency deployment and monitoring.
- Techniques are developed for cross-market comparison of energy-savings potential.
- We find significant cost-effective energy-savings potential from efficiency for Indian and Chinese refrigerators.

## ARTICLE INFO

*Keywords:*
Energy efficiency
Appliances
Web crawling
Data mining

## ABSTRACT

Appliance energy-efficiency programs are a central component of many countries' energy-policy portfolios. A major barrier to optimal implementation of these programs is lack of data to determine market baselines, assess the potential for cost-effective energy savings, and track markets over time to evaluate and verify program impacts. To address this gap, we have developed the International Database of Efficient Appliances (IDEA), a suite of software tools that automatically gathers data that is currently dispersed across various online sources and compiles it into a unified repository of information on efficiency, price, and features for a diversity of appliances and devices in markets around the world. In this article we describe the framework and functionality of IDEA, and we demonstrate its power as a resource for research and policy development related to appliance energy efficiency. Using IDEA data for refrigerators in China and India, we assess the potential for cost-effective energy savings within each market by computing robust indicators that can also be easily compared across different appliances and markets. We find that significant cost-effective savings are available on both markets. We discuss implications for the development of future energy-efficiency deployment programs.

## 1. Introduction

Improving energy efficiency (EE) has long been recognized as an effective means of reducing energy demand, often at a cost that is below the marginal price of energy [1]. To meet their Nationally Determined Contributions under the Paris climate agreement, many nations plan to accelerate the deployment of EE for appliances, via policy actions such as mandatory minimum energy performance standards (MEPS) or appliance labeling programs that aid consumer decision making by providing information on energy consumption. Globally, the most common standards and labeling (S & L) programs use the categorical-labeling approach, in which governments define, for each covered product, a number of efficiency levels (ELs) starting from a MEPS level, and require that a label be displayed at the point of sale indicating each product's EL. In recent years, expansion of appliance S & L programs has been the focus of a broad multi-national effort through programs such as the Clean Energy Ministerial's Super-Efficient Appliance Deployment (SEAD) initiative [2], the United Nations Environment Programme's United for Efficiency initiative [3], as well as a variety of non-governmental-organization (NGO)-supported efforts in individual emerging economies. Among the barriers to advancement of these programs is a lack of appropriate market data to support policy development.

Effective program development requires an accurate understanding of the mix of product efficiencies available on the market as well as the relations among price, EE, and other product features. Such information is needed for assessments of the available savings potential on a given

* Corresponding author.
*E-mail address:* bfgerke@lbl.gov (B.F. Gerke).
[1] Current address: Energy Solutions, 449 15th St. #400, Oakland, CA 94612, USA.

http://dx.doi.org/10.1016/j.apenergy.2017.07.093
Received 31 January 2017; Received in revised form 28 April 2017; Accepted 23 July 2017
0306-2619/ 

market, for analyses of the consumer costs and benefits associated with proposed EE policies and programs, and for *ex-ante* estimates of key program impacts, such as national energy savings or consumer financial savings. Moreover, data that can track average market efficiency levels over time are important for *ex-post* measurement and verification of actual program impacts.

Analytical inputs for S & L policy development have traditionally been developed using bottom-up engineering analyses that estimate manufacturing costs at different EE levels, coupled with a series of assumed markups along the supply chain to yield a final relation between EE and consumer price (e.g., [4–6]). Although this approach is useful for analyzing the underlying costs of EE improvements in the manufacturing sector, it does not account for the complex market forces that shape real-word prices. Indeed, past research has shown that real appliance prices have undergone a long-term decline [7,8], consistent with the effects of technological learning in the manufacture of appliances. As a result, traditional cost-benefit analyses, using constant price estimates, have overestimated the actual costs of MEPS in the United States [9]. Improved analyses that incorporate price trends can produce better cost estimates and, in some cases, justify more stringent EE policies [10]. Recent research has also shown evidence for an interesting dynamic in which appliance EE policies themselves may help to drive down the price of efficient technologies [11–13], possibly by eliminating manufacturer incentives to bundle EE with other high-end product features. Thus, to develop and maintain optimal appliance EE policies, it is important to consider detailed market data that can reveal the actual on-market relations among price, EE, and product features, as well as their evolution over time.

Historically, however, such data have been fragmentary, expensive to obtain, or completely unavailable in most markets. Price information resides mainly with retailers or market research firms, complete product feature information resides with manufacturers, and EE data are either stored in certification databases maintained by government or industry groups or are available inconsistently if at all from manufacturers. Many of these data are closely held by firms and other entities that have a vested interest in EE policy decisions. Thus, when relevant data are made available to policy makers, the time horizon covered by the data may be sharply limited, and the accuracy of the data can be difficult to verify. The ready availability of online information creates an opportunity to improve the availability of relevant information, using automated software tools to collect, organize, and manage the data. The Billion Prices Project [14] at the Massachusetts Institute of Technology has demonstrated the viability of online data as a resource for developing high-frequency economic price indices and performing detailed economic research on product prices, including the impacts of economic policy [15]. Related research has recently shown that online prices correspond closely with prices found in physical stores, across a range of global markets [16], strengthening the case for online data as a tool for market analysis.

To leverage the internet data resource in the service of appliance EE deployment, we have developed the International Database of Efficient Appliances (IDEA). IDEA is the first complete implementation of the SEAD data-access framework [17,18], which defines the structure of a unified database of global appliance information. IDEA draws 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, which are currently dispersed across the internet, combining the information from them into a rich data set containing information on appliance price, efficiency, and features. This data set has many valuable applications in EE policy development and evaluation. The automated data-collection tools allow regular data updates, which enables real-time tracking of market trends, including improvements in EE performance within a given market or price declines for new, high-efficiency technologies, such as solid-state lighting products [19]. Furthermore, IDEA's global scope facilitates international EE benchmarking, and it can provide insights into the international impacts of national EE policies, such as cross-market spillover or dumping of inefficient products in unregulated markets.

A functioning implementation of IDEA is complete and is collecting data for several appliances and countries. The resulting data have been used in two previously published reports [20,21]. This paper offers a comprehensive overview of the IDEA framework, functionality, and target applications. We also demonstrate the potential of IDEA to produce insights that can inform EE S & L policy development. Specifically, we analyze data on refrigerators in the Chinese and Indian markets that was compiled in IDEA by drawing from various online retailers and from each country's S & L certification database. We perform a set of regression analyses to determine the relation among price, EE, and various other product features. From the resulting models, we develop EE cost-effectiveness indicators that enable direct comparison of EE savings potential among different markets. In this example analysis, we find that, as of mid-2015, significant potential existed for cost-effective energy savings from EE in both China's and India's refrigerator markets.

Section 2 of this paper describes IDEA in detail, including example applications and strategies for data collection and management. Section 3 describes the data used in this paper and lays out methods for two example analyses. Section 4 presents the results of these analyses, including market snapshot analyses of refrigerator EE in India and China as well as an approach to assessing and cross-comparing the potential for cost-effective refrigerator EE upgrades in both countries. In Section 5, we briefly discuss potential implications for the relevant S & L programs as well as the broader potential of IDEA for global appliance EE deployment.

## 2. The database: applications and functionality

IDEA is envisioned as a comprehensive database of information on EE, price, and features of appliances for sale across a broad range of international markets. The database contains detailed information on the individual appliance models being sold on each market covered, including the following types of data:

- Identifying information (e.g., model numbers, brands, and unique product identifiers such as universal product codes [UPCs] or international article numbers [EANs]).
- Price data, including both current prices and an archive of historical prices observed for each product.
- Product features and attributes.
- Data on product availability and popularity.
- EE data certified by governments or other reputable sources.
- Region-specific data, such as currency, units of measure, and particular EE metrics and test procedures.

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 software-based data-collection techniques, as discussed below. Hereafter, we will refer to these data generally as *retail data*. Where possible, retail data are imported directly from public-facing *application programmer interfaces* (APIs), such as the Best Buy Developer Portal in the U.S. [22]. These interfaces allow programmatic access to the data underlying individual product listings on a website. When APIs are not available or are not sufficiently comprehensive, data can be collected using web crawlers—i.e., custom-built computer programs that systematically visit the product listings on a retail or manufacturer site and parse them to extract the details for each product. In principle, the IDEA framework could also incorporate retail data collected by visiting physical retail stores in each country, but this approach entails high costs and logistical difficulty, and the data are more difficult to validate. Therefore, IDEA focuses on data collected from online retailers, with the goal of creating a database that represents the internet market in each country.