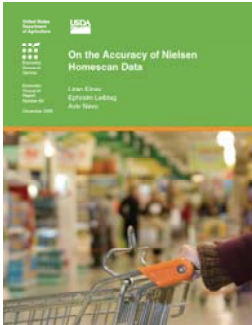


# ERS *Report Summary*

Economic Research Service

December 2008

U.S. Department of Agriculture



This is a summary  
of an ERS report.

Find the full report at  
[www.ers.usda.gov/  
publications/err69](http://www.ers.usda.gov/publications/err69)

## On the Accuracy of Nielsen Homescan Data

**Liran Einav, Ephraim Leibtag, and Aviv Nevo**

Researchers use Nielsen Homescan data, which provide detailed food-purchase information from a panel of U.S. households, to study the dynamics of retail food markets.

### **What Is the Issue?**

Some questions have been raised regarding the credibility of the Nielsen Homescan data because the data are self-recorded and the recording process is time-consuming. Given the time commitment, households who agree to participate in the sample might not be representative of the U.S. population as a whole, and those who agree to participate may not record their purchases accurately.

### **What Did the Study Find?**

The analysis conducted in this report suggests that the Homescan data contain recording errors in several dimensions, but that the overall accuracy of self-reported data by Homescan panelists seems to be in line with other commonly used (government-collected) economic data sets.

For approximately 20 percent of food-shopping trips recorded in the Nielsen Homescan data, there was no corresponding transaction in the retailer's data, suggesting that either the store or date information was recorded with error. Using the retailer's loyalty card information, the study finds some shopping trips that did not match up with Nielsen Homescan data, implying that households did not record all of their trips in their Homescan records.

For the trips that did match up, roughly 20 percent of the items purchased were not recorded. For those items that were recorded, quantity was reported fairly accurately: 94 percent of the quantity information matched in the two data sets. The match for prices was lower: in almost half of the cases, the two data sets did not agree. However, much of this difference can be attributed to transactions that involved promotional or other temporary sale prices in either the Nielsen Homescan data or the retailer's data.

Nielsen's practice of using store-level data as an estimate of what households actually paid, poses a challenge when those stores have multiple possible prices in a given time period due to loyalty card or other shopper-specific price promotions. Indeed, for prices that involve no promotion or temporary price reduction, there are recording errors in only about 17 percent of the cases. Therefore, much of the overall price difference is likely caused by the way Nielsen imputes prices and not by recording errors by the panelists. Mismatched prices would most likely be less of a problem for stores that only have one price per product in a given week, so that the results highlight the importance of store pricing practices in food price analysis.

ERS is a primary source of economic research and analysis from the U.S. Department of Agriculture, providing timely information on economic and policy issues related to agriculture, food, the environment, and rural America.

[www.ers.usda.gov](http://www.ers.usda.gov)

The study also compares the recording errors to errors in other commonly used economic data sets, and finds that errors in Homescan are of the same order of magnitude, for example, as reporting errors in earnings and employment-status data.

### ***How Was the Study Conducted?***

Homescan records contain all products purchased by a household on a particular day in a particular store, as they were scanned by the consumer. The study compared these records to data obtained from a single retailer. The retailer's data contain the products purchased in each of the transactions at the same store and day reported by the household, as recorded by the cashier. Using data from trips made during 2004, the records from both data sets were matched. The matched transactions were compared and contrasted, and differences in various dimensions were recorded. In order to study the impact the recording errors might make in an applied study, the price paid was regressed on household characteristics in both data sets to see if the results differed.