Taking Some of the Uncertainty Out of Retailer Self-Reported ENERGY STAR® Market Share

Can We At Least Believe in the Trends?

Dr. Harley H. Barnes, Aspen Systems Corporation, Rockville, MD

ABSTRACT

ENERGY STAR market share is a key indicator of consumer preference for ENERGY STAR-qualified products. As such, it is a key metric for market-transformation program evaluation. However, difficulty in acquiring actual ENERGY STAR market shares and doubts about how accurately retailers are willing to report their true sales data cloud its value for use in evaluations.

The research reported in this paper addresses this issue. It reduces some of the uncertainty in the use of retailer self-reported ENERGY STAR market shares for residential appliance products by correlating trends in self-reported market share with trends in an independently measured metric that serves as a plausible surrogate for ENERGY STAR market share. The surrogate variable is the proportion of appliance models on display in retailers' showrooms that are ENERGY STAR models.

Three different evaluation metrics for market-share trend and display-proportion trend are developed, correlated, and tested for four common residential appliances (twelve tests). Eleven of the twelve correlations are found to be statistically significant. The paper concludes that trends in retailer self-reported ENERGY STAR market share are valid indicators of trends in actual ENERGY STAR market share. As such, retailer self-reported ENERGY STAR market share has important value for assessing progress in ENERGY STAR market-transformation programs.

The Issue

ENERGY STAR market share is one of the key metrics for the evaluation of market-transformation program progress. It is a strong indicator of one of the major goals of these programs: increased consumer preference for ENERGY STAR-qualified products.[1]

In spite of the key importance of ENERGY STAR market share for evaluating the progress of market-transformation programs, however, existing methods for collecting *actual* sales data to measure it (principally point-of-sale scanned data) do not fully satisfy the need. This is because these methods (1) produce market-share data that either are not geographically congruent with program territories; or (2) omit key segments of the market such as consumers who shop in stores that will not provide scanned point-of-sale sales data; or (3) require that retailers grant access to their actual sales data for tabulation.

To help compensate for these deficiencies, at least five major regions of the country[2] and the U.S. Department of Energy and U.S. Environmental Protection Agency are using local retailer *self-*

[1] An Association of Energy Service Professionals Brown Bag Seminar[SM] recently emphasized the importance of market share as an indicator of market-transformation program progress. (AESP 2003). "ENERGY STAR market share," as used in this paper, means the proportion of all sales of a product type during a given period of time, such as a month, that are ENERGY STAR-qualified.

reported market share as evidence of progress in appliance market-transformation programs. Yet, self-reported ENERGY STAR market share also has deficiencies as a market-transformation indicator. This is because (1) retailers resist submitting the data, and (2) the accuracy of the self-reported data is uncertain. Further, for reasons of practicality and cost, the accuracy of self-reported sales data can almost never be verified against actual sales data. Therefore, substantial uncertainty surrounds the use of self-reported sales data for evaluation purposes.

The research reported in this paper addresses the latter issue. It proposes to reduce some of the uncertainty in the use of self-reported market share as a metric for evaluation purposes. The paper will provide statistical evidence that the *trend* in retailer self-reported ENERGY STAR appliance market-share data can be used as a valid indicator of the actual *trend* in ENERGY STAR market share.[3]

The data used in the research are derived from data collected to monitor the progress of New York State's ENERGY STAR Products Program (Products Program), one of the **New York Energy $mart^{SM}** portfolio of programs. The Products Program is a market transformation program that promotes increased consumer preferences for ENERGY STAR residential appliances.[4]

Research Design

Retailer self-reported ENERGY STAR market share can be accepted, or rejected, as a valid indicator of actual ENERGY STAR market-share *trend* if a hypothesis that expresses the proposition can be formed, operationalized, and tested. A source of independent measurements is available that can be used for this purpose, provided certain assumptions can be accepted. This section will:

- Describe these assumptions
- Develop the testable hypothesis
- Describe the data available to test the hypothesis, and
- Describe the test statistic for the hypothesis.[5]

Assumptions and Development of a Testable Hypothesis

The paper's research approach and testable hypotheses depend entirely on the following hypothesis, which will be assumed to be true:

H_1: Retailers allocate different product models, including ENERGY STAR models, to their showroom in the proportions that they expect customers to buy, or in the proportions they intend to encourage them to buy. The hypothesis pertains to a specified period of time, e.g., a month.

The following scenarios illustrate this hypothesis. If retailers expect 50% of their customers will be willing to buy ENERGY STAR-qualified room air conditioners during the summer, they will stock and display 50% ENERGY STAR-qualified room air conditioners for the summer season. Or if they have a reason to try to achieve a 50% ENERGY STAR market share during a particular time period, as

[2] The states and utilities participating in the Northwest Energy Efficiency Alliance, New York, Wisconsin, California, and the states and utilities participating in the New England Energy Efficiency Partnership.
[3] The paper limits itself to investigation of this claim; it does not address the issue of whether observed trends may be attributed to specific programs.
[4] See Barnes 2005 for a summary description of NYSERDA's Products Program.
[5] For readers who want to skip the technical details of the research, the results of the hypothesis test and this paper are presented beginning with the section titled, "Results: Reducing the Uncertainty in Retailer Self-reported Sales Data."

is the case with performance-based trade promotions, they will stock and display at least 50% ENERGY STAR-qualified products of the category affected by the promotion. In this discussion, the proportion of all models of a product on display that are ENERGY STAR qualified models will be called the product's "ENERGY STAR display proportion."

Many retailers participating in NYSERDA's Products Program have stated this "rule" during the program's annual interviews with managers of participating stores and in retailer focus groups (NYSERDA 1999-2004, NYSERDA 2000). This is not to say that this is the only reason retail store managers have given for their stocking and ENERGY STAR display proportions. Other popular rationales include, "the corporation makes the choice of product mix for us," or "we stock what the buyers' group offers." It seems highly reasonable, however, that these higher-level decision makers operate by the same rule as the more independent store managers.

The assumption of truth for H_1, i.e., the claim that sales are related to what is on display, may seem to be a stretch. If ENERGY STAR market share for appliances could be measured directly, such assumptions would not be needed. However, it cannot be measured directly without great difficulty, and therefore the validity of retailer self-reported market share remains an important issue for market-transformation evaluation. If a relationship between retailer self-reported ENERGY STAR market share and a plausible, independently measured surrogate such as ENERGY STAR display proportion can help to reduce the uncertainty in self-reported market share, it is worth exploring.

H_1 can be restated in an equivalent form that is more relevant for this research:

H_1': The proportion of actual showroom sales of an appliance category during a given time period that is ENERGY STAR-qualified (actual ENERGY STAR market share) is related to the proportion of all models of that appliance category on display during the same period that is ENERGY STAR-qualified (ENERGY STAR display proportion).

The null version of this hypothesis is testable using the data available for this paper, but it is a hypothesis about cross-sectional relationships during a specific period of time. That is, the hypothesis is about the proportion of sales in a specific month that are ENERGY STAR qualified and the proportion of models on display that are ENERGY STAR qualified during the *same* month. It is not as interesting for evaluation purposes as a hypothesis that extends it to *consecutive periods of time*. That is, if H_1' is true for one period of time, then it should follow that the actual ENERGY STAR market shares and ENERGY STAR display proportions will continue to be related over consecutive time periods. This statement about *trend* relationships is the more interesting hypothesis for evaluation purposes and the one that this research tested.[6]

It follows from this that another hypothesis needs to be stated and accepted as an assumption: H_1 and H_1' are independent of time. Every year since NYSERDA began interviewing store managers in 1999, retailers have given the "rule" stated by H_1 as one of the reasons for their stocking choices. Therefore, the following hypothesis is accepted without further discussion:

H_2: The stocking principle stated in H_1 does not change over time. Retailers follow it for every new product order.

[6] Although H_1' is the predecessor for a hypothesis about trend relationships, it does not have to be separately tested in order to test whether the relationship exists over consecutive periods of time. Market-transformation program evaluators are interested in what is happening to *markets* over time. For this purpose, they aggregate outcomes such as store market shares, for specific periods of time, e.g., years, into indicators. Then they compare the indicators from year to year. The relationships in these longitudinal time series can be examined without examining the cross-sectional relationships that give rise to them.

H_2 allows us to compare trend data without worrying about whether some extraneous influence has caused H_1 and H_1' to become invalid over time and interfere with our test of trend relationships.

This leads to the hypothesis that is the focus of this research.

H_3: The time trend in retailers' self-reported ENERGY STAR market shares is related to the time trend in ENERGY STAR display proportion in the retailers' showrooms.

H_3 cannot be tested because we do not have a quantitative value for the hypothesized relationship; however, we can test the null version of this hypothesis:

H_0: The time trend in retailers' self-reported ENERGY STAR market shares is not related to (is independent of) the time trend in ENERGY STAR display proportion in the retailers' showrooms.

The two time-trend variables that are the subject of H_0 can be defined in terms of interval data or ordinal (ranked) data. This research intentionally avoids any assumptions about the accuracy of retailers' interval-level self-reports. That would stretch credibility too far. Further, as will be discussed later, there are theoretical reasons for avoiding such assumptions. This research investigates the trend relationships in the rank, or relative, order of the data. That is the same as asking, as retailer self-reported ENERGY STAR market share changes, does the ENERGY STAR display proportion change in the same *direction*? This becomes important later for the selection of a test statistic.

If the research can reject H_0, evaluators will have empirical support for believing that a trend in retailer *self-reported* ENERGY STAR market share serves as a valid indicator of the trend in *actual* ENERGY STAR market share.

The Research Data for Appliance Market Shares and ENERGY STAR display proportions

NYSERDA's Products Program has been measuring the proportions of ENERGY STAR appliance models on display in participating retailers' showrooms since 1999. By 2004, the Products Program had made seven measurements on this variable. Four appliance categories have a full time series of seven measurements: refrigerators, dishwashers, clothes washers, and room air conditioners. The measurements are made by NYSERDA's implementation contractor and are independent of retailer self-reports of sales data.

Figure 1 shows the six-year trends (seven store surveys) for ENERGY STAR display proportions.[7] The ENERGY STAR display proportion values in Figure 1 are the survey-by-survey simple averages of the measured store ENERGY STAR display proportions. The values of *n* shown are the number of appliance stores surveyed. They do not necessarily indicate the number of stores selling each product surveyed.

[7] All data used in this research have been developed from these surveys and tabulations of participating retailer self-reported sales data prepared by Aspen Systems Corporation for NYSERDA.

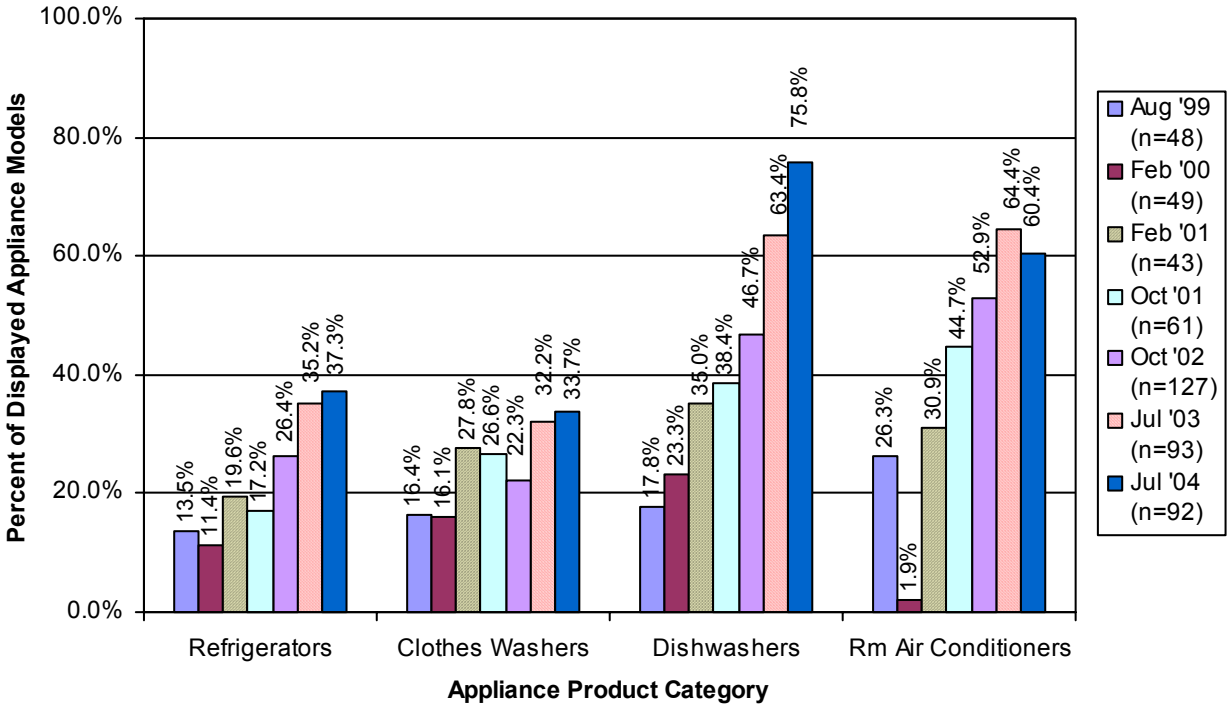

Figure 1. Trends in the Average Store ENERGY STAR Display Proportions for Appliance Models, Products Program Participating Showrooms, Aug 1999 – Jul 2004

Concurrently, NYSERDA has asked participating retailers to submit ENERGY STAR and total sales data voluntarily every month. The number of appliance retailers submitting sales data has grown over the years from 65 in 1999 to over 260 in 2004. The average proportion of participating retailers submitting sales data has varied from 22% in 1999 to 94% in 2004.[8] These participating-retailer numbers do not include partners in the national ENERGY STAR program that submit data to DOE, e.g., Sears, Lowe's.

The sales data used in this research exclude sales over the Internet and sales to home-remodeling contractors because these data are not related to showroom display choices. They also exclude sales and ENERGY STAR display-proportions data from retailers that are partners of the national ENERGY STAR program.

The ENERGY STAR display proportions and market shares were measured for the same month, but a difference exists in the time of the month that they represent. ENERGY STAR display proportion was measured on the day the store was surveyed. The market shares are for the entire month in which the survey was conducted. This introduces an unknown degree of uncertainty into the results; however, in view of the hypothesis and the robustness of the methodology used to test it (to be described), this should not have a major effect on the findings.

Figure 2 shows the trends in the average of the retailers' self-reported ENERGY STAR market shares for the months during which the display-area proportions in Figure 1 were measured.

[8] The reporting count and percentage for 1999 are averages across the four appliance product categories in the month of August, the only month that year in which these data were recorded. The count and percentage for 2004 are the averages of the monthly reporting data across the four appliances for the entire year.

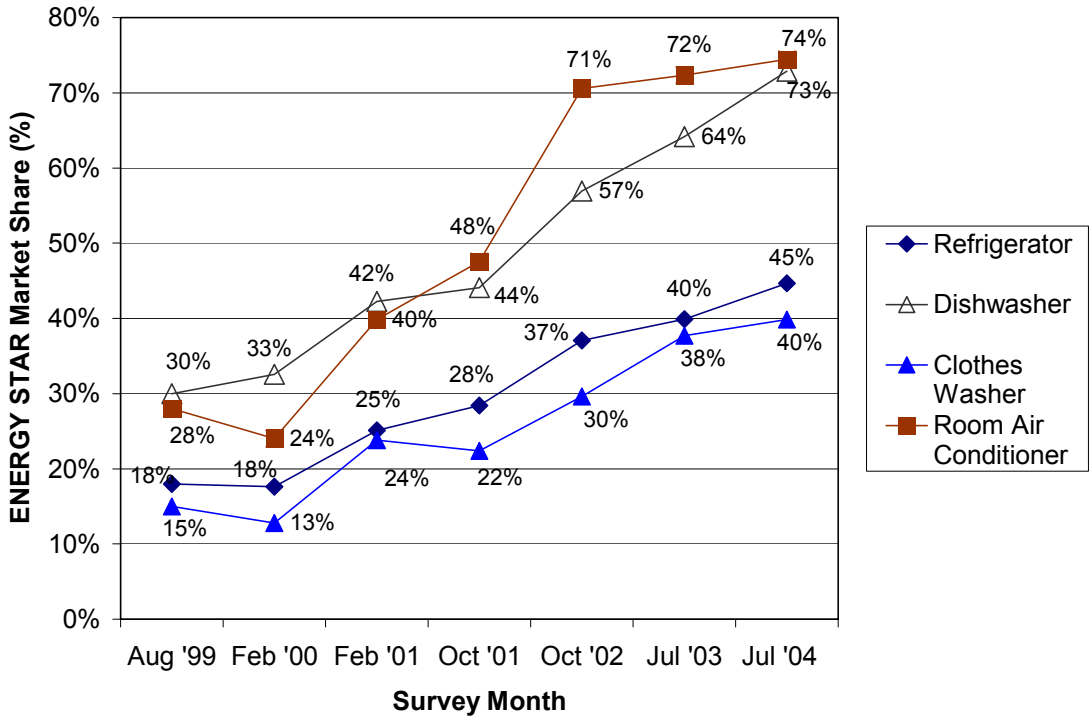

Figure 2. Trends in Self-Reported ENERGY STAR Appliance Market Shares for the Months in Which the ENERGY STAR Display Proportions Were Measured, Participating Showrooms, Aug 1999 – Jul 2004

The evaluation metrics used for the hypothesis test (to be described in the next section) required that the ENERGY STAR display-proportion and market-share data used in each time period be from same stores. Inasmuch as the stores from which the two measurements came were chosen by different methods, and neither was a census (at least not in the early years of sales-data collection), this necessarily resulted in the loss of data when the two sets of data were merged. Table 1 shows the number of stores for which both display-proportion and market-share data were available for the same store in each of the seven surveys. These are the stores that supplied the research data.

Table 1. Number of Stores Having Both Display-Proportion and Market-Share Data for Each Survey (*n* for the Evaluation Metrics)

Store Survey Month	Refrigerators	Dishwashers	Clothes Washers	Room Air Conditioners
August 1999	3	3	3	1
February 2000	14	15	15	11
February 2001	23	23	23	10
October 2001	48	45	47	34
October 2003	62	57	63	69
July 2003	55	53	55	58
July 2004	67	61	62	67

Consideration was given to dropping the August 1999 data from the research, but ultimately they were retained. On inspection, the ENERGY STAR display proportions and market shares from that early period were reasonable relative to values from stores without matching data in that period;

therefore, they provided at least an indication of the 1999 ENERGY STAR display proportions and market shares. For ranking purposes, this would be adequate.

Selecting Evaluation Metrics for the Hypothesis Test

Evaluators will aggregate raw display and sales data to create evaluation metrics that satisfy their objectives. They have several options for such aggregation. For example, Figures 1 and 2 used the simple average of the sampled stores' ENERGY STAR market shares and display proportions, but the raw data might also have been aggregated to calculate the ENERGY STAR market shares and display proportions without considering what the individual stores were doing. The latter would consist of the ratio of the aggregated ENERGY STAR sales to all aggregated sales and the ratio of aggregated ENERGY STAR models on display to all aggregated models on display. The author is aware of efficiency-program metrics in use that illustrate both approaches. The research reported in this paper will be more useful to evaluators if it tests H_0 using alternative metrics such as these. Table 2 shows five ways the data available for this paper might be aggregated to track a program's progress.

Table 2. Five Evaluation Metrics That Can Be Created from Raw ENERGY STAR Display-Proportion and Market-Share Trend Data

Evaluation Metric	How Calculated
ENERGY STAR Display Proportion	
(D1) Estimated simple average of store ENERGY STAR display proportions. (Sample)	Measure the store ENERGY STAR display proportions and compute their simple average. (Sample)
(D2) Estimated ratio of all ENERGY STAR models on display to all models on display. (Sample)	Count ENERGY STAR models on display across all sampled stores and divide by the count of all models on display in these stores (in effect, a weighted store average). (Sample)
ENERGY STAR Market Share	
(M1) Estimated simple average of store ENERGY STAR market shares. (Sample)	Measure the ENERGY STAR market share for each individual store in the display-proportions survey during the survey month and compute their simple average. (Sample)
(M2) Estimated ratio of all ENERGY STAR unit sales to all unit sales. (Sample)	Count the ENERGY STAR unit sales by stores in the display-proportions survey during the survey month and divide by the count of all unit sales by these stores (in effect, a weighted store average). (Sample)
(M3) Ratio of all ENERGY STAR unit sales to all unit sales for <u>all</u> reporting stores. (All available sales data)	Count ENERGY STAR unit sales in all reporting stores (not just the stores in the display survey) during the survey month and divide by the count of all unit sales in these stores (in effect, a weighted store average). (All available data)

Recognizing that evaluators have options such as these, this research tested H_0 on the ENERGY STAR display-proportion and market-share relationships created by three combinations of the metrics listed in Table 2. The three combinations selected were chosen because (1) they test H_0 using the metrics most likely to be encountered in evaluation practice, and (2) with one exception, they provide symmetry in the method by which the market-share and display-proportion variables are calculated.

Table 3 shows the combinations tested. Test #1 (a combination of D1 and M1) tests H_0 using simple-average store metrics. Test #2 (a combination of D1 and M3—the one exception) tests H_0 using a mix of the metrics that is in actual use by a program with which the author is associated. Test #3 (a combination of D2 and M2) tests H_0 using the aggregated-data-ratio metrics.

Table 3. The Three Combinations of Metrics Tested in the Research

Test No.	ENERGY STAR Display Proportion Metric	ENERGY STAR Market Share Metric
1.	(D1) Estimated simple average of store ENERGY STAR display proportions. (Sampled stores)	(M1) Estimated simple average of store ENERGY STAR market shares. (Sampled stores)
2.	(D1) Estimated simple average of store ENERGY STAR display proportions. (Sampled stores)	(M3) Ratio of all ENERGY STAR unit sales to all unit sales for all reporting stores. (All available store data)
3.	(D2) Estimated ratio of all ENERGY STAR models on display to all models on display. (Sampled stores)	(M2) Estimated ratio of all ENERGY STAR unit sales to all unit sales. (Sampled stores)

The Test Statistic

Correlation methods that rely on interval data are inappropriate as test statistics for several reasons. First, retailers' self-reported sales data should not be taken as measurements of true sales; therefore, the true interval differences between periodic market-share reports cannot be known with confidence.[9] Second, market-share trends are, theoretically, not linear. They follow an "S" or bell-shaped curve, depending on how they are defined.[10] And third, interval-level trend data beg to be projected, and such projections would be inappropriate for the first and second reasons.

Correlation methods that use rank-ordered data provide statistics that can be used to test for a relationship between the directions in which two variables are trending. Such a method was selected for testing the relationship between the ENERGY STAR market-share and display-proportion trend data. For this research each of the interval-level metrics listed in Table 3 were calculated and ranked for each variable. This resulted in seven pairs of rank-ordered data for testing. Table 4 illustrates the results of this process for the refrigerator metrics and test no. 1.

Table 4. Illustration of How the Refrigerator Metrics Were Ranked and Paired for Test #1

Survey	Average Store Market Share (M1)	Average Store ENERGY STAR Display Proportion (D1)	Ranks for the Average Store Market Shares	Ranks for the Average Store ENERGY STAR Display Proportions
Aug '99	0.167	0.177	1	4
Feb '00	0.176	0.051	2	1
Feb '01	0.229	0.137	3	2
Oct '01	0.347	0.170	4	3
Oct '02	0.375	0.216	5	5
July '03	0.395	0.305	7	6
Jul '04	0.382	0.327	6	7

[9] Many retailers do not maintain electronic records of ENERGY STAR sales. The author's experience indicates that such retailers will avoid the burden of having to identify ENERGY STAR models manually from their sales records. They will "estimate" the proportion of their sales that they think were ENERGY STAR-qualified (their own estimate of ENERGY STAR market share) and calculate their ENERGY STAR sales using that proportion. In these cases, the accuracy of the ENERGY STAR sales data depends on the retail managers' knowledge of ENERGY STAR models and skill in guessing the proportion of sales that consisted of those models. Thus, for many retailers, the true interval differences between their sequential reported market shares cannot be known with confidence.

[10] Market-share trends may be assumed to be linear for periods of time such as the six years covered by the data for this research. An early test of whether this research might succeed was performed using linear regression. However, for the longer term, linear models will not be theoretically applicable.

The rank-order correlation test statistic selected is Spearman's *r* (r_s), also known in some texts as Spearman's ρ (rho). It was evaluated against one other rank-order statistic (Kendall's tau) and selected for its relative simplicity of calculation and ease of interpretation.[11] Spearman's *r* is the application of the product-moment correlation to rank data. Like the product-moment correlation, when the ranks are in perfect positive agreement, $r_s = +1.0$, and when they are in perfect inverse agreement, $r_s = -1.0$.

Two factors influenced the choice of a critical region for the test of significance. First, the ENERGY STAR display proportions were measured on different days during the survey month because the surveyors could not be in every sampled store on the same day. The sales data for calculating market share for individual stores, on the other hand, are reported for the entire month that included the day on which the store was surveyed. If the surveyor measured ENERGY STAR display proportion near the end of the month, after most of the ENERGY STAR models had been sold, or perhaps, just after the store manager had restocked the showroom, the correspondence between the two indicators would not be optimal. Therefore, perfect agreement in the ranks was not expected.

Second, examination of Figures 1 and 2, and the assumption contained in H_1, indicate that there is a tendency for larger values of ENERGY STAR display proportion to be paired with larger values of market share. This suggested that a one-tailed test should be used to test for significance.

The imperfect timing of the match-ups of the two variables, the use of a one-tailed test, and the wish to have convincing results were all balanced to choose a one-tailed critical region (α) of 0.05 for the test of significance for each correlation. The next section presents the results of the twelve tests (three combinations of evaluation metrics for each of the four appliance categories).

Results: Reducing the Uncertainty in Retailer Self-reported Sales Data

Table 5 shows the results of the three tests of the null hypothesis (no relationship between the trend in retailer self-reported ENERGY STAR market share and the trend in ENERGY STAR display proportion). The values of r_s indicate the observed strength of the relationship between these two trends for each of the four appliance categories.

Table 5. Results of the Spearman's *r* Tests

Product Category	Spearman's *r* for a Positive Relationship between Trends in Retailer Self-reported ENERGY STAR Market Share and ENERGY STAR Display Proportion (Significant values of r_s for $\alpha = 0.05$ are **bold.**[12]) ($n = 7$ for all tests)		
	Test No. 1 (Evaluation Metrics D1 paired with M1)	Test No. 2 (Evaluation Metrics D1 paired with M3)	Test No. 3 (Evaluation Metrics D2 paired with M2)
Refrigerators	**0.750**	**0.964**	0.536
Dishwashers	**0.857**	**1.000**	**0.750**
Clothes Washers	**0.714**	**0.893**	**0.821**
Room Air Conditioners	**1.000**	**0.964**	**0.964**

[11] Explanations of Spearman's *r* and other rank-order correlation methods may by found in Blalock 1972, Section 18.4, and Conover 1980, Section 5.4.
[12] Conover 1982, Appendix A10.

All of the rank-order correlations are significant except that for refrigerators under test number 3. Test number 3 correlated the estimated ENERGY STAR market-share ratios calculated for each survey by dividing the aggregated ENERGY STAR sales by the total sales and the estimated ENERGY STAR display proportions calculated by dividing the aggregated ENERGY STAR model-display counts by the total model-display counts.

Conclusions & Applications: The Value of Retailer Self-reported Sales Data

Table 5 indicates that a statistically significant positive relationship exists between the trend in retailer self-reported ENERGY STAR market share and the trend in ENERGY STAR display proportion for each of the three sets of evaluation metrics for dishwashers, clothes washers, and room air conditioners. The comparable results for refrigerators are generally consistent, but not as strongly in the predicted direction.

From the number of significant correlations, the null hypothesis is rejected and it is concluded that a trend in retailer self-reported market share is a valid indicator of the actual ENERGY STAR market-share trend (H_3).

If we are willing to accept hypothesis H_1' about the relationship of ENERGY STAR display proportion to actual ENERGY STAR market share, these results indicate that market-transformation programs can validly use trends in retailer self-reported ENERGY STAR appliance market-share to track progress in changing consumer preferences for ENERGY STAR appliances.

The application of this finding can be illustrated. This paper began by claiming that ENERGY STAR market share is one of the key metrics for the evaluation of market-transformation program progress. Market-transformation evaluators face the following questions at four important stages of program operation:

1. For a start-up program: Has consumer preference for ENERGY STAR products begun increasing yet?
2. For a program that has already demonstrated progress in growing consumer preferences: Is consumer preference for ENERGY STAR products still increasing?
3. For a program that has demonstrated progress in growing consumer preferences and now may have indication that it has reached a market-share ceiling: Has a maximum level of consumer preference for ENERGY STAR products (relative to the current program design) been achieved?
4. For a terminated program or a reduced level of program intervention: Is the achieved level of consumer preference for ENERGY STAR products being sustained?

All of these questions can be answered by trend data. If retailer self-reported market-share trend data are valid, then retailer self-reported market share has considerable merit for evaluating the progress of market-transformation programs.

A paradox exists with respect to the third and fourth questions. When a retailer self-reported appliance market-share trend flattens out for several years, a repeated rank-order correlation test of data such as that used in this research may indicate that no relationship exists between the two indicators. Yet the fact that a significant relationship was found in their trends as they were increasing suggests that the flat market-share trend is also a valid trend. In the opinion of the author, this paradox adds to the usefulness of retailer self-reported market share as an indicator of market-transformation progress. A flattened market-share trend would indicate (1) that market share has reached its maximum and it is time to transition the transformation program to a sustaining program (or roll out a more aggressive

marketing strategy), and (2) if the flat trend persists, or decreases slightly and settles out, that the transformation has been sustained.

Acknowledgements

I am grateful to the following individuals at Aspen Systems Corporation who assisted at different stages in the research for this paper: Nancy Hassett, Blanche Swaim, Jeanne Liu, Mark Quigley, and Rosa Cassidy.

References

Association of Energy Services Professionals (AESP). 2003. "Tracking Market Penetration." AESP Brown Bag Seminar[SM], June 3.

Barnes, H. 2005. "Market Transformation: Half a Decade of Results from a Supply-side Intervention." Proceedings of the 2005 International Energy Program Evaluation Conference, Brooklyn, N.Y., August 17 - 19. In press.

Blalock, H. 1972. *Social Statistics*, 2nd Edition. New York: McGraw-Hill.

Conover, W.J. 1980. *Practical Nonparametric Statistics,* 2nd Edition. New York: John Wiley & Sons.

New York State Energy Research and Development Authority (NYSERDA). 1999-2004. "Final Baseline Report," December 29, 2000. "New York ENERGY STAR® Products Program Participant Practices Report," dated April 13, 2000; August 9, 2001; February 15, 2002; December 13, 2002; December 30, 2003; and December 22, 2004 (in draft). All reports prepared by Aspen Systems Corporation.

New York State Energy Research and Development Authority (NYSERDA). 2000. "Report of Focus Groups with Retailers and Multifamily Property Managers." December 20. Prepared by Aspen Systems Corporation.