The Valuation of Energy Efficiency in Homes

Prepared for:
U.S. Department of Housing and Urban Development
Washington, D.C.

Prepared by:
Michael Carliner, Liza Bowles, and Joe Nebbia
Newport Partners LLC

December 2008

Acknowledgements

Special acknowledgement goes to Dr. Regina Gray, Policy Development and Research, U.S. Department of Housing and Urban Development, for her overall project guidance.

We thank Paul Emrath, National Association of Home Builders, and Daran Wastchak, D.R. Wastchak, LLC for participating in a panel at the EEBA Excellence and Building Conference and Expo where the initial results of this project were reported. We are especially grateful to all HERS providers who contributed vital data for this study, as well as Abt Associates for their various work supporting this project.

Advisory Panel

Newport assembled an advisory panel of experts to help guide the development and implementation of the project, and to point Newport in the direction of resources they would need. The panel was comprised of members from industry, government and academia including Steve Baden, Residential Energy Services Network (RESNET); Paul Bishop, National Association of Realtors (NAR); Don Boucher, The Appraisal Institute; Dr. Paul Emrath, National Association of Homebuilders (NAHB); Nick Gromicko, International Association of Certified Home Inspectors; Dr. Joe Laquatra, Cornell University; Sam Rashkin, US EPA Energy Star Homes; and Dr. Kevin Simmons, Austin College. This group of experts provided advice on economic analysis of the housing markets; insight into how information on a home was transferred during a sale; and guidance on the availability of data for the project. They provided an assessment of the feasibility of the proposed repeat sales index methodology and gave advice on adjustments. The panel also reviewed the content of the survey and gave suggestions for edits. Two members of the panel signed a letter of introduction to HERS providers asking for their help in the repeat sales index. The Advisory Panel held two official conference calls, and members communicated individually with Newport throughout the project. One member sat on a panel where the preliminary results were discussed at the Energy and Environmental Building Association's (EEBA) annual conference. Another member delegated his panel position to another member in his organization.

Table of Contents

Executive Summary

The value of energy efficiency in houses is a difficult thing to measure, but of vital importance to the efforts of energy efficiency advocates. If energy efficiency is a valuable feature in a home that is considered desirable by both builders and buyers, and if houses with these features are able to hold that value, energy efficiency becomes a much more marketable product or strategy. Both hedonic regressions and willingness-to-pay studies have been conducted examining the value of energy efficient technology. However, studies have not established the difference in appreciation rates between homes that have energy efficient features and homes built using standard building practices.

This study is Phase 1 of what is designed to be a multiphase project with this first phase investigating and establishing a methodology. The original Phase 1 study design called for in-depth case studies to examine this question from a qualitative standpoint. After determining that data availability and the cost-prohibitive nature of this approach made the methodology unrealistic, Newport worked with the Advisory Committee to determine an alternate approach. After considering various approaches, the most promising was the creation of a repeat sales index that compared the appreciation rates of energy efficient homes against the appreciation rates of standard homes. Because specific energy efficient features of a home are hard to isolate, Newport decided to focus on homes certified under an energy efficiency program. This approach used Energy Star certified homes as the energy efficient homes and baseline homes in the same Metropolitan Statistical Area as the control group. Using this method, Newport examined 1,748 Energy Star properties and 13,845 control group properties in the Dallas/Ft. Worth/Arlington, TX Metropolitan Statistical Area (MSA); as well as 643 Energy Star properties and 1,701 control group properties in the Columbus, OH MSA. Some of these properties were sold more than twice. In the Dallas/Ft. Worth/Arlington MSA, appreciation rates of Energy Star homes were slightly higher than the control group. In the Columbus MSA, appreciation rates of Energy Star homes were slightly lower. In both cases, there was no statistical significance to the difference in appreciation rates.

If Energy Star homes have the same appreciation rate as control groups, this means that they maintain the premium originally paid for the home. If this is the case, buyers could purchase Energy Star homes with confidence that they would enjoy the benefits while they lived there and would not lose this premium at resale. If Energy Star homes appreciate at a higher rate than control groups, energy efficiency would be an even better investment. If they appreciate at lower rates than control groups, it is a signal to the energy efficiency community that marketing and communication of value needs to be improved.

Newport also created a survey that could be used to collect data on how information is communicated on the energy efficiency of that home, and who is communicating that information. The survey was designed to shed light on how information flows about energy efficiency in home sales – if at all – and how this process could be improved. The survey was not implemented due to time and funding constraints; however, it would be a valuable research tool to implement in Phase 2.

Phase 1 succeeded in establishing a methodology for how future studies could be conducted. If a larger dataset, made up of energy efficient and control homes in at least 20 MSAs, was available, a significant answer might be possible on a national scale. A larger dataset would need to focus on MSAs in a variety of climates and with a variety of market conditions. Buy-in from the Energy Star community and the RESNET community as a whole – or alternatively, a different energy efficiency program – will be necessary in order to collect data for, and to conduct, a broader study. This buy-in would allow researchers to choose the most appropriate MSAs for further analysis. Further details on the energy efficient features of homes – such as whether insulation or mechanical equipment was being used to increase efficiency – would also be helpful in order to help explain the results of such a study. Some energy efficient features in houses are "invisible." It may not be obvious to a consumer that a home has a high level of insulation, or that the house has low air leakage because of good sealing techniques. Other features, such as appliances or air conditioners may need to be replaced sooner than something such as insulation. A broader study – covering a longer period of time in order to include economic highs, lows, and transitions – that is based on the repeat sales index used in this study, with buy-in from industry groups holding important data, would be a repeatable way to measure appreciation of energy efficient homes compared to standard homes.

In summary, this study discovered that:

- Using a repeat sales index as a way to compare the appreciation rates of energy efficient and standard homes is a valid model for relating home price appreciation to energy efficiency in homes.
- The study showed no difference between the appreciation rates of Energy Star homes compared to standard homes in two MSAs analyzed.
- A broader study over a longer time period, covering a larger number of MSAs is needed in order to generalize results across the country. The study, using the methodology described in this report, could more firmly establish whether or not energy efficient homes have different appreciation rates than standard homes.
- Industry buy-in is important for any study of Energy Star homes focusing on multiple MSAs across the country.

Introduction

When purchasing a home, a prospective buyer examines the market and makes a decision based on the aspects of a home that they value, the priorities they set on these aspects, and their ability to pay for the features that they desire. A buyer's decisions show their willingness to pay for certain features of a home. However, are the buyers aware of the energy efficient features of a home? If they are aware, is the next buyer aware when the home is resold? Does the home appreciate in a way that it keeps the premium originally paid for energy efficient features over standard homes? If not, does the premium increase or decrease? These questions point to the important issue of how energy efficiency is valued in the home, how that value is communicated, and whether homes retain that value over time. This subject is vital to the efforts of energy efficiency advocates. If energy efficient features of a home hold or increase their value, it makes them an easier sell to consumers. Energy Efficient features of a home that retain their value or increase over time would make more buyers willing to pay the premium, knowing that they would recoup their cost upon resale. However, if the energy efficient features of homes decrease in value over time, buyers may behave differently.

There could be many explanations for why a home would or would not retain the premium paid originally for energy efficient features. The actual appreciation rates of energy efficient homes, and those of standard homes, are needed to determine whether or not this premium is retained. Information on how energy efficient features are advertised and who among the parties involved in a home sale (seller, buyer, real estate agent, appraiser, lender, etc.) is aware of these features and makes them a priority during the selling process would help clarify why appreciation rates behave the way they do.

This report details:

- initial attempts to approach these issues through detailed case-studies examining how information on energy efficient homes is communicated (found in Appendix B);
- a repeat sales index measuring the appreciation of Energy Star homes compared to standard homes;
- the need for a survey examining what buyers value in a home, and how information about the home is communicated; and ,
- the results of the repeat sales index and implications for the energy efficient homes market.

Background

Beyond Willingness-to-Pay

Although research exists describing the association between energy efficient technology and home value, the methodologies used in that research have been incomplete, or have not answered the question about how energy efficient homes appreciate. The most direct research (Nevin and Watson, 1998) used hedonic regression analysis on 1991-1996 American Housing Survey data. Those results suggest that house values are $20 higher for each $1 reduction in utility costs. More recent articles (Black, 2003, 2004) have cited the 20:1 ratio, using it in payback calculations, but have not included resale value when estimating an 11% return on investments in solar electric systems. Hedonic regressions are beneficial for analyzing a point in time, but the ultimate test of energy savings being capitalized into the value of a house requires time series data. Cost effectiveness studies can predict the utility savings over the life of the energy system, but they cannot determine market acceptance as revealed by resale values. According to Dacquisto, *et al.*, the hedonic regression studies that have been done do not "produce a single result, or range of results, for the implicit price associated with a measure of energy efficiency that could be effectively used by builders or consumers to impute value, or by appraisers to adjust the prices of otherwise comparable housing units." (Dacquisto, *et al.*, 2001)

There have also been numerous willingness-to-pay studies that ask home owners the amount they are willing to pay for certain features of a home. In contrast to the hedonic estimates such as those from Nevin and Watson, surveys asking home buyers about their willingness to pay for energy efficiency have not shown that they are willing to pay a substantial premium. For example, in a survey of recent and prospective homebuyers conducted by the National Association of Home Builders in 2007, one of the questions was: "How much would you pay up front, in the purchase price of your next home, if it would save $1,000 every year in utility costs?" The median response was $5,000 (Ahluwalia, 2008), implying an after-tax required rate of return of 20 percent. Other surveys have produced similar results. Moreover, there is generally a bias in consumer surveys toward ***overstating*** willingness to pay, as it is easier to spend hypothetical money than to actually come up with the funds.

It is not clear why consumers are not willing to pay a larger premium for energy efficiency, but one reason may be that they don't expect to live in the house indefinitely, and they are afraid that they won't be able to recapture the premium upon resale.

If energy efficiency improvements consist of structural features that can be expected to permanently reduce energy use, then rising energy prices may be expected to lead to higher appreciation. Some vendors assert that such above-average appreciation can be expected, without any empirical basis. If consumers are concerned about merely recouping their investments, however, then a demonstration that appreciation is not lower for energy-efficient homes should have a positive impact on demand. Hedonic regression and willingness-to-pay studies do not provide insight into the long-term value of energy efficient features in a house investment. In order to determine the investment value of the energy efficient technology, this study looks beyond the willingness-to-pay and hedonic regression numbers to find the actual appreciation difference between energy efficient homes and non-energy efficient homes, if such a difference exists.

Defining Perception as a Barrier to Energy Efficiency in the Housing Market

In order for innovative technology to gain recognition as a value-booster in homes, acceptance must come at all levels of the housing market. Designers, builders, homebuyers, realtors, appraisers, and lenders all play a role in determining how highly specific aspects of homes are valued (as well as how this value is communicated) – including innovative technologies in the home.

- Designers and builders make market decisions on what materials and technologies to use in the home that will bring the best value in the original sale.
- Consumers determine how much they are willing to pay for innovative technologies in a home. This decision may be based on how much the technology will save the buyer through energy savings or durability; or on the appeal or comfort of the new technology; or finally, on the marginal increase to the resale value of the home.
- Real estate agents advise buyers and sellers on the value of their home, and how to best market the home's qualities.
- Appraisers estimate the market value of a home and its components. Their estimates directly influence and inform a lender's decision on how much they are willing to invest in a home when making mortgage risk assessments.

All of these players work together to determine what aspects of a home add to its market value, and to what extent they cause the home to appreciate. The perception of whether or not having a specific technology in a home, or energy efficiency in general, will add to the resale value on a home goes far in determining the real value that these technologies will add to a home. If a consumer believes that a home with a certain technology, or energy efficiency in general, will have a higher resale value, then they are more likely to pay a higher price for the home. If consumers consistently pay higher prices for homes with these technologies, appraisers may put more value in the technology component, realtors may market more heavily toward innovative houses, and lenders may be more willing to invest at higher levels in technologically advanced houses. In order to avoid the barrier of perception, comprehensive data are needed to determine the increased resale value of homes as a result of innovative technologies in the home.

Additional information is also needed to determine the impacts of information exchange on the process and how they influence the valuation of energy efficiency in housing.

Study

Repeat Sales Index

Trying to measure house prices, and the effect of characteristics such as energy efficiency on market values, is beset by several complications. There are so many ways in which homes may differ in terms of structural characteristics and condition, as well as location, that it is virtually impossible to specify all the parameters of difference. Home sales data available for statistical analysis rarely include information about more than a few characteristics, such as number of bedrooms and bathrooms and square feet of living area.

For measuring the impact of a characteristic such as energy efficiency on market value at a given point in time, taking into account other characteristics that affect value, the typical practice is to estimate a hedonic model. Sales prices for a sample of homes are compared (regressed) against characteristics of the homes, with the regression coefficients for each of the characteristics representing implicit prices. This requires information about many characteristics and assumptions about how the characteristics interact. The implicit price estimates may be biased because some important characteristics are not included (and are correlated with characteristics that are included) or because the relationship is misspecified (e.g., joint effects are assumed to be additive when they are multiplicative).

In addition to the complexity of the bundle of characteristics incorporated in homes and heterogeneity in the housing stock, measurement of house price appreciation is complicated by the fact that homes are sold infrequently, at irregular intervals, so we cannot observe the prices of two groups of homes at two or more specific times and compare the changes.

The hedonic method could be used to estimate changes in the implicit prices for each characteristic over time, and the effects of those changes in values for characteristics could be applied to a "standard" house with a fixed set of characteristics, but that still requires more information about characteristics than is generally available. That approach is used by the Census Bureau to construct an index of new home prices.

In this report, an alternative method, repeat sales regression, is used. The repeat sales method takes information from sales of homes that were sold twice or more in an area. Sales prices at different times for the same house are compared, so the characteristics will be the same, unless the structure has been altered between sales. Even though sales occur at irregular intervals, estimates of the change in value during intervening, regular, periods are produced.

In this report, repeat sales estimates of appreciation for homes with and without energy ratings are compared. Even though prices for the same houses within each group are compared, differences in relative rates of appreciation could be due to changes in the value associated with other characteristics, rather than the one of interest. Fireplaces or decks, for example, may become more or less fashionable, and may be concentrated in one of the groups. To minimize those problems, therefore, it was important to compare the "treatment" group to a "control" group that was as similar as possible in terms of characteristics such as size, year built, location, etc.

If house i is sold for price $P_{i,t}$ at time t and sold again for price $P_{i,t+n}$ at time t+n, we can describe the relative price as

$$\frac{P_{i,t+n}}{P_{i,t}} = \frac{I_{t+n}}{I_t} \cdot X_i$$

where I_t and I_{t+n} are values for an index of prices for all homes in the sample at times t and t+n, and X_i is the appreciation for the individual house i during that period compared to the appreciation for the index. The objective is to find values for the index that explain as much as possible of the changes observed for all homes and attribute as little as possible to idiosyncratic changes for individual homes.

The calculation is simplified by taking logs of both sides of equation 1:

$$\text{Log}\,(P_{i,t+n}) - \text{Log}\,(P_{i,t}) = \text{Log}\,(I_{t+n}) - \text{Log}\,(I_t) + \text{Log}\,(X_i)$$

The estimation procedure finds values ($\hat{I}_t$) for the price index in each period that minimize the sum of the squares of $\log(x_i)$.

An alternative way of looking at the change in the price of house *i* between *t* and *t+n* is to consider the appreciation over that range as reflecting the cumulative effect of changes in the market in each intervening period, along with the idiosyncratic change:

$$\text{Log}\,(P_{i,t+n}) - \text{Log}\,(P_{i,t}) = \text{Log}\,(R_{t+1}) + \text{Log}\,(R_{t+2}) + \cdots + \text{Log}\,(R_{t+n}) + \text{Log}\,(X_i)$$

where R_{t+k} is appreciation for the group in the single period from *t+k-1* to *t+k*. It turns out that this calculation produces the same result as the estimate of the index. That is,

$$\hat{I}_{t+n} = I_t \times \hat{R}_{t+1} \times \hat{R}_{t+2} \times \cdots \times \hat{R}_{t+n}$$

Where $\hat{I}$ and $\hat{R}$ represent the estimated values.

The index value for the first period is not estimated, but is set equal to the base value (e.g. 100). Whether we calculate an index directly or calculate period relatives, each sale pair affects not only the estimates for the periods in which sales occurred, but also the estimates for intervening periods.

Figures 1 and 2 illustrate the type of information used to construct a repeat sales index.

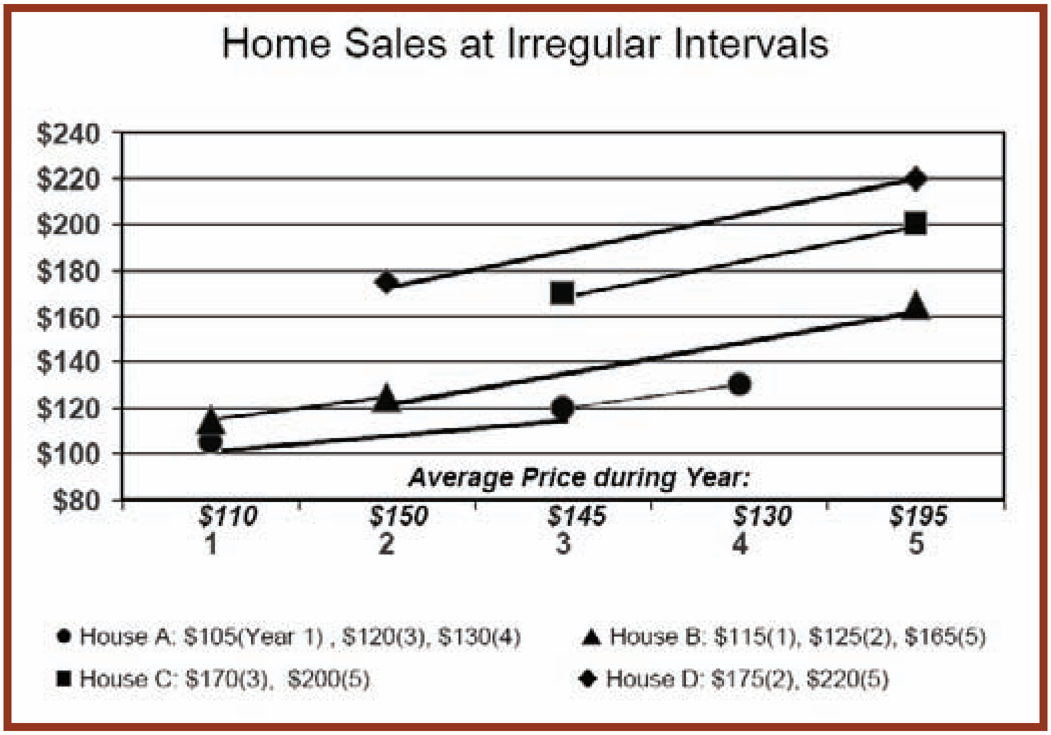


Figure 1

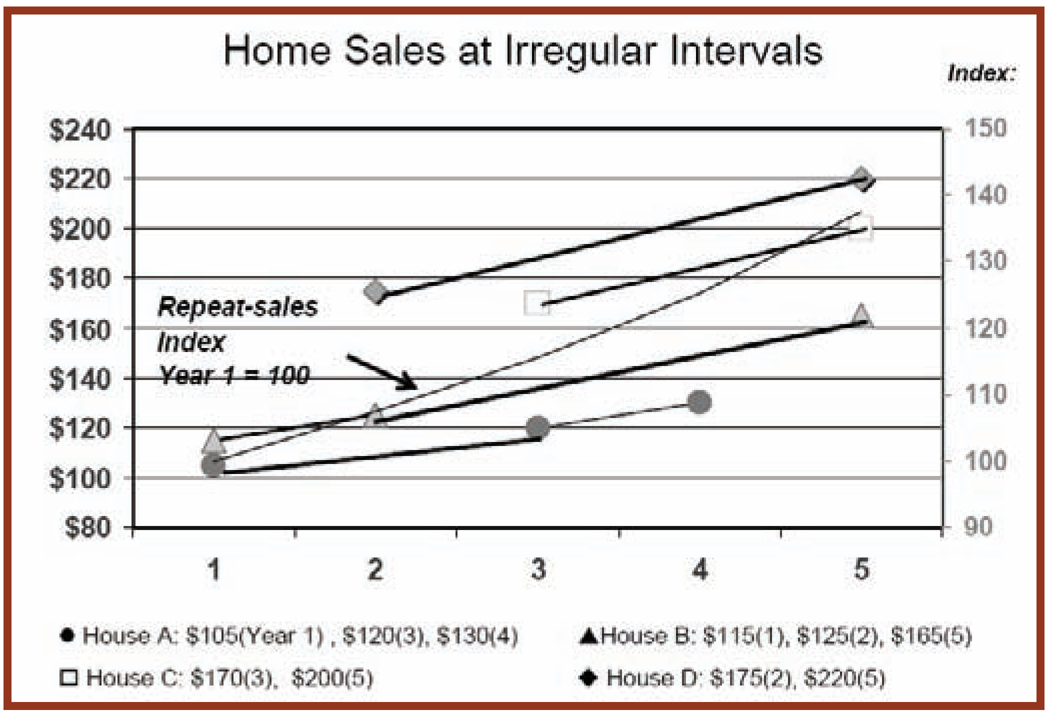


Figure 2

Here we have 4 houses observed over 5 periods. Houses A and B are each sold three times, so each provides 2 pairs of successive transactions. Houses C and D are each sold twice. All together, we have 6 transaction pairs. Notice that all 4 houses in this hypothetical example have upward price trajectories, but if we were to measure price trends by the average price for sales in each period, it would appear that prices declined in period 3 and 4, because of the mix of homes that were sold.

To calculate the repeat sales index, we create a variable representing each period (id1...id5). If the first sale in a transaction pair (two sales of a property) is in that period, the value of the period variable for that transaction pair equals -1. If the second sale in a sales pair is in that period, the period variable equals +1. Otherwise, for that observation (transaction pair) the period variable equals zero. The data and regression results are shown in Table 1.

Table 1

Repeat Sales Example

Pair/Observation	1	2	3	4	5	6
House	A	A	B	B	C	D
Year – 1st Sale in Pair	1	3	1	2	3	2
Year – 2nd Sale in Pair	3	4	2	5	5	5
Price 1st Sale in Pair	$105	$120	$115	$125	$170	$175
Price 2nd Sale in Pair	$120	$130	$125	$165	$200	$220
Relative	1.14	1.08	1.09	1.32	1.18	1.26
ln(rel)	0.1335	0.0800	0.0834	0.2776	0.1625	0.2288
Index Dummies:						
id1	-1	0	-1	0	0	0
id2	0	0	1	-1	0	-1
id3	1	-1	0	0	-1	0
id4	0	1	0	0	0	0
id5	0	0	0	1	1	1
Relative Dummies:						
rd1	0	0	0	0	0	0
rd2	1	0	1	0	0	0
rd3	1	0	0	1	0	1
rd4	0	1	0	1	1	1
rd5	0	0	0	1	1	1

Index Regression

. regress logrel id2 id3 id4 id5 , noconst
Number of obs = 6
F(4, 2) = 55.82
Prob > F = 0.0177
R-squared = 0.9911
Adj R-squared = 0.9734
Root MSE = .02881

Var	Coef	S.E.	t-stat	*Index*
id2	**0.072**	**0.024**	**2.95**	*107.44*
id3	**0.145**	**0.024**	**5.96**	*115.62*
id4	**0.225**	**0.038**	**5.97**	*125.25*
id5	**0.319**	**0.027**	**11.97**	*137.61*

Period Relative Regression

. regress logrel rd2 rd3 rd4 rd5, noconst
Number of obs = 6
F(4, 2) = 55.82
Prob > F = 0.0177
R-squared = 0.9911
Adj R-squared = 0.9734
Root MSE = .02881

Var	Coef	S.E.	t-stat	*Relative*
rd2	**0.072**	**0.024**	**2.95**	*1.0744*
rd3	**0.073**	**0.027**	**2.75**	*1.0761*
rd4	**0.080**	**0.029**	**2.78**	*1.0833*
rd5	**0.094**	**0.038**	**2.49**	*1.0986*

In the estimation process, the cross-sectional dummy variable for the first period is excluded, effectively forcing its coefficient to be equal to zero. Since the dependent variable is in terms of logs, the antilog of the coefficient for that quarter = e^0 = 1.00, and that base period value is multiplied by 100. Index values for other quarters are calculated similarly, based on the estimated coefficients. For example, the coefficient for the second period is 0.072, so the index value was 100 x $e^{0.072}$ = 107.44.

To calculate period relatives, the dummy variables equal 1 in each period after the first sale in a transaction pair until the second sale in a transaction pair and zero otherwise (rd1..rd5). The coefficients from the period relative regression represent logs of ratios of prices in the current period to prices in the previous period, estimated so as to minimize the relative deviations for individual sales pairs.

In this example, the calculated index explains nearly all the changes in prices for individual homes, reflected in the high R-squared values. In the real world, prices for individual houses don't move together in such a neat fashion.

Background on Repeat Sales Method:

The regression-based repeat sales method for constructing price indices was developed by Bailey, Muth, and Norse (1963). It was generally not utilized for another quarter-century, however, in part because of the absence of appropriate software, including address-matching routines, as well as the absence of data in convenient form. Interest in the technique was revived by Case and Shiller (1987), and it received a push upon its adoption by the Office of Federal Housing Enterprise Oversight (OFHEO) as an element in the capital stress test used in the regulation of Fannie Mae and Freddie Mac mandated by the Federal Housing Enterprises Safety and Soundness Act of 1992. OFHEO was incorporated into the Federal Housing Finance Agency (FHFA) under the Housing and Economic recovery Act of 2008.

There are now several repeat sales indices for metropolitan areas, states, and the U.S. The most widely known and used repeat sales measures are those from FHFA/OFHEO. The FHFA/OFHEO indices are based on information from mortgages purchased and/or securitized by Fannie Mae and Freddie Mac. A similar set of indices, calculated slightly differently but based on the same data, are reported by Freddie Mac.

Another well-established set of regularly reported indices were developed by Case and Shiller (CS), and are now produced by Standard and Poor's. Specific CS indices are the basis for futures contracts traded on the Chicago Mercantile Exchange. The CS data are not limited to homes with mortgages that passed to Fannie Mae and Freddie Mac, but instead are based on public transaction records assembled by private vendors, similar to the data used in this study.

There are some complex econometric issues connected to the calculation of repeat sales indices. One issue that has gotten particular attention is whether the variance of the individual error terms ($\log(x_i)$ in terms of the specification above) is related to the time between transactions. Most of the general repeat sales estimates, such as those from Case-Shiller and OFHEO, use weighted least squares (WLS), with lower weights given to sales pairs that are further apart. For such WLS calculations, the first step is to calculate indices in the manner described here using ordinary least squares (OLS). Squares of the residuals from the first-stage regression are then regressed against the number of periods between sales, (and possibly the square of the number of periods as well). Using the resulting estimated relationship of sample variance to time between sales to determine appropriate weights, the index regressions are then recalculated. This procedure was tried with the data in this study, and, although there was some relationship between the variance of the residuals and the number of quarters between sales, recalculation of the indices using weighted least squares had virtually no effect on the results. For simplicity, this report only uses the OLS results. If we had data covering a longer time span, rather than only the period from 2002 to 2007, the effects might have been greater.

Data Requirements for this Study:

To use the repeat sales index as an indication of relative appreciation for this study Newport required two sets of data:

- A group of energy efficient homes – Energy Star homes for this study

- A group of houses in the same area as the Energy Star homes, and built around the same time, but built using standard building practices and features.

The actual numbers being compared between the datasets were the appreciation rates as determined by sales amounts found in property transaction data. In order to isolate the effect of energy efficient features as much as possible, Newport needed to ensure that houses in similar markets and of similar size were being compared. The two datasets had the following requirements:

- The Energy Star homes and control group homes needed to come from the same MSA
- The Energy Star homes and control group homes needed to be built around the same time
- The Energy Star homes and control group homes needed to be of comparable sizes.

In order to determine the appreciation rates of homes in the study, as well as whether or not homes were built around the same time and of comparable sizes, Newport required a source of data that would provide:

- The full address of a property
- Prices for all sales of a property
- Sale dates for all sales of a property
- Square footage
- Number of bedrooms and bathrooms
- Year built

Additionally, the study required the identification of the addresses of Energy Star homes for which transaction records would be needed. After collection of the data, the two datasets had to be combined so that Energy Star addresses could be matched with transaction records and non-Energy Star addresses could be identified within the transaction data to act as the control group.

Why Energy Star for Homes?

The repeat sales index needed a group of energy efficient homes to compare to a control group of homes without the energy features. The study chose to use only Energy Star homes as the energy efficient group for several reasons:

1. Energy Star program claims to have certified 750,000 homes as of 2006 providing a large sample with which to work.
2. As a national program, it allowed more possibility of eventually conducting a nationwide study than other regional programs. (See Energy Star website on milestones)
3. Although, within any program, there will be great variation in the actual technologies and strategies used, as well as the extent of energy efficiency achieved above minimum requirements, some variability is limited by studying only one program. For example, using both Energy Star homes and a green building program adds different levels of efficiency that will vary from group to group. An Energy Star home might not have the same green technologies. The green home may have a higher or lower energy efficiency requirement than Energy Star. Within one program, appreciation rates may still vary based on whether energy efficiency is achieved through longer lasting features such as insulation, or more visible features such as heat pumps. However, using one program allows for the most similarity across the sample.
4. Because Energy Star for Homes started in 1995, there was a longer amount of time that could be studied under this program than under many other existing programs.
5. All HERS providers are required to keep electronic records of all homes certified under Energy Star, and therefore, would be a good source of data.
6. Two members of the advisory panel were closely connected with the Energy Star for Homes program and were valuable sources of information for how to go about creating the study and finding data.

Energy Star Record Keeping:

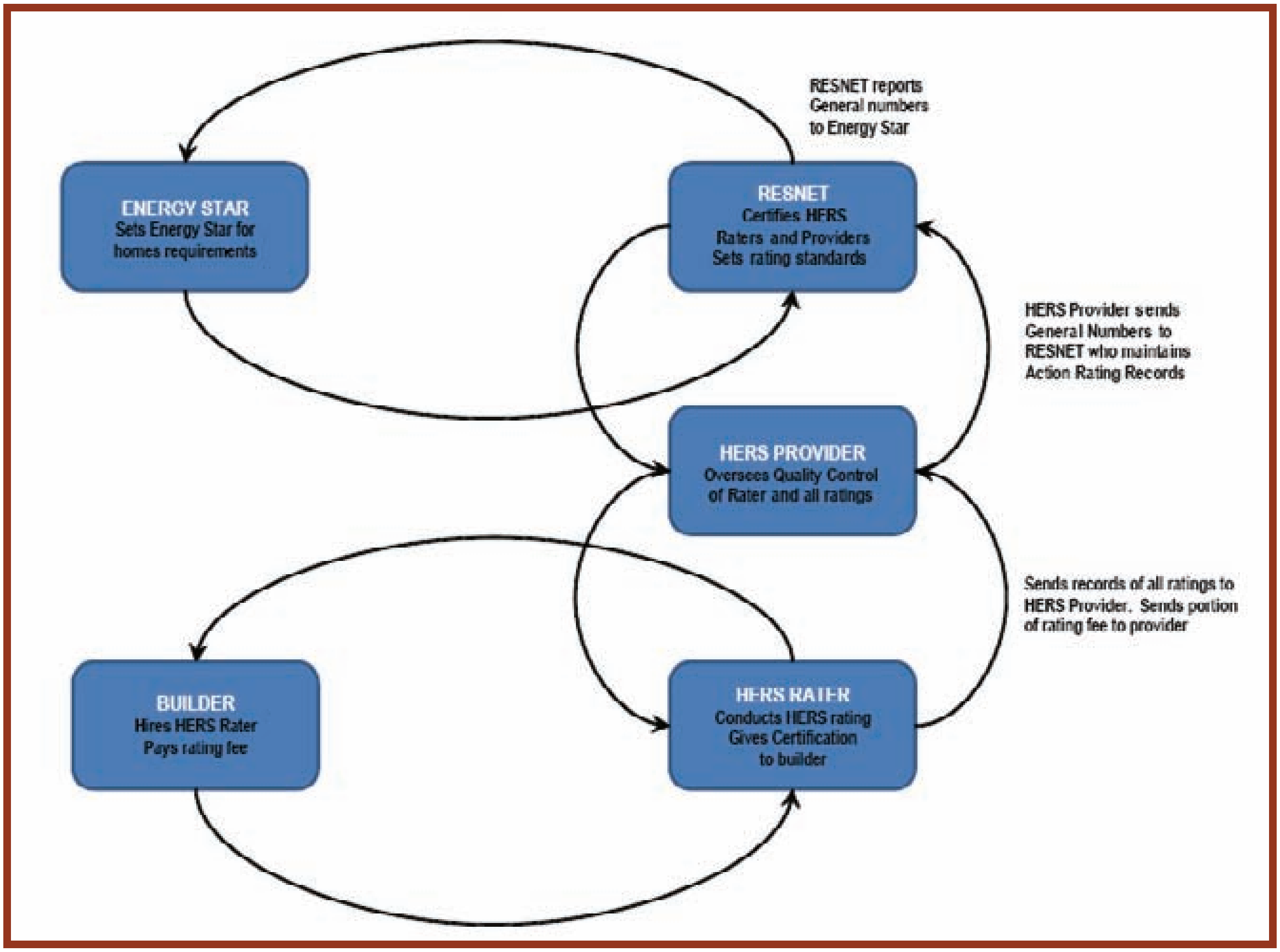
In trying to obtain records of Energy Star addresses, Newport discovered the following record keeping structure within the Energy Star program (see Table 2):

Table 2

Environmental Protection Agency and the US Department of Energy	EPA and US DOE do not maintain records of homes rated and certified under the Energy Star program. EPA tracks the total number of homes but does not collect specific information for each home.
Residential Energy Services Network (RESNET)	Energy Star relies on RESNET to maintain all records of houses rated and certified under the Energy Star for Homes program. RESNET is the organization that developed the HERS rating system and that certifies HERS raters and HERS providers that manage the rating for Energy Star homes.
HERS provider	Every HERS rater is connected to a HERS provider who is responsible for quality assurance and administrative aspects of HERS ratings. According to RESNET Standards, all HERS providers must "maintain an electronic database of information for each home rated." (See RESNET Standard 2006)
HERS rater	A RESNET certified rater conducts the HERS rating of an Energy Star Home, and documents this rating.

Figure 3 gives a visual representation of how these relationships work:

Figure 3

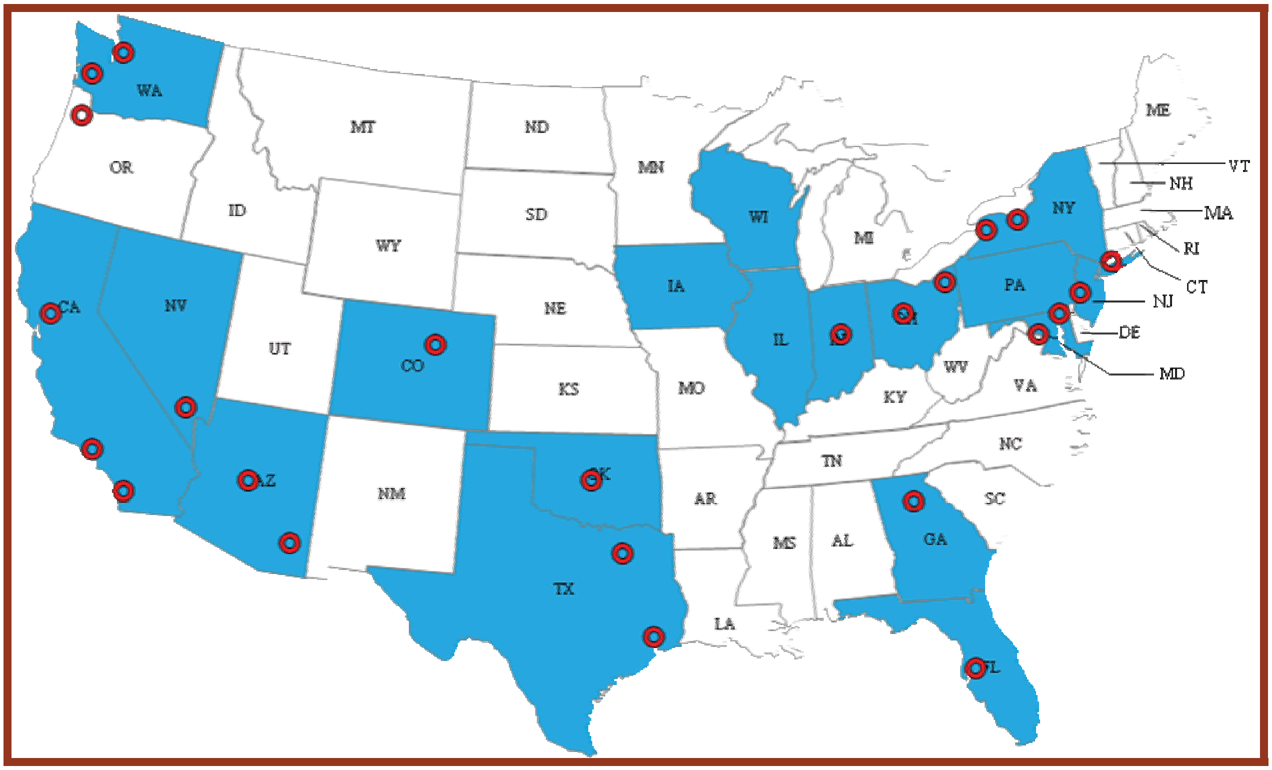


HERS providers therefore, were chosen as the target data source for identifying Energy Star homes as they would have records of all homes certified by raters working under them. RESNET is currently working to develop a national database that will house all HERS rating records and will be passed from raters, through HERS providers, to RESNET. This database is not in existence now, but would make future studies on this topic much easier.

Outreach to HERS Providers:

The search for HERS providers that could serve as data sources for a repeat sales index began by identifying the states with the most Energy Star homes. Within those states, the Metropolitan Statistical Areas (MSAs) with the highest concentration of Energy Star homes were identified. (See Energy Star for Homes website) Figure 4 shows the states with high Energy Star numbers highlighted in Blue, and the MSA's in which HERS providers were contacted marked with a red circle. In many cases, several HERS providers were contacted within a single MSA.

Figure 4



HERS providers in targeted MSAs received an initial introduction letter via email explaining the project and how they could help by proving records of all Energy Star homes for which they acted as HERS provider. In addition to being signed by Newport, the letter also contained endorsement signatures from the Department of Housing and Urban Development, the Environmental Protection Agency's Energy Star for Homes program, and the Residential Energy Services Network (RESNET). Several HERS providers that spoke with us about the project and agreed to participate mentioned those endorsements as a key reason that they were willing to send records. A generic copy of this letter can be found in Appendix A. Email letters were followed up with phone calls and further emails. Although it was difficult to raise interest among them, HERS providers with records in 8 MSAs expressed interest in participating. Actual participation only came from HERS providers with records in the MSAs of:

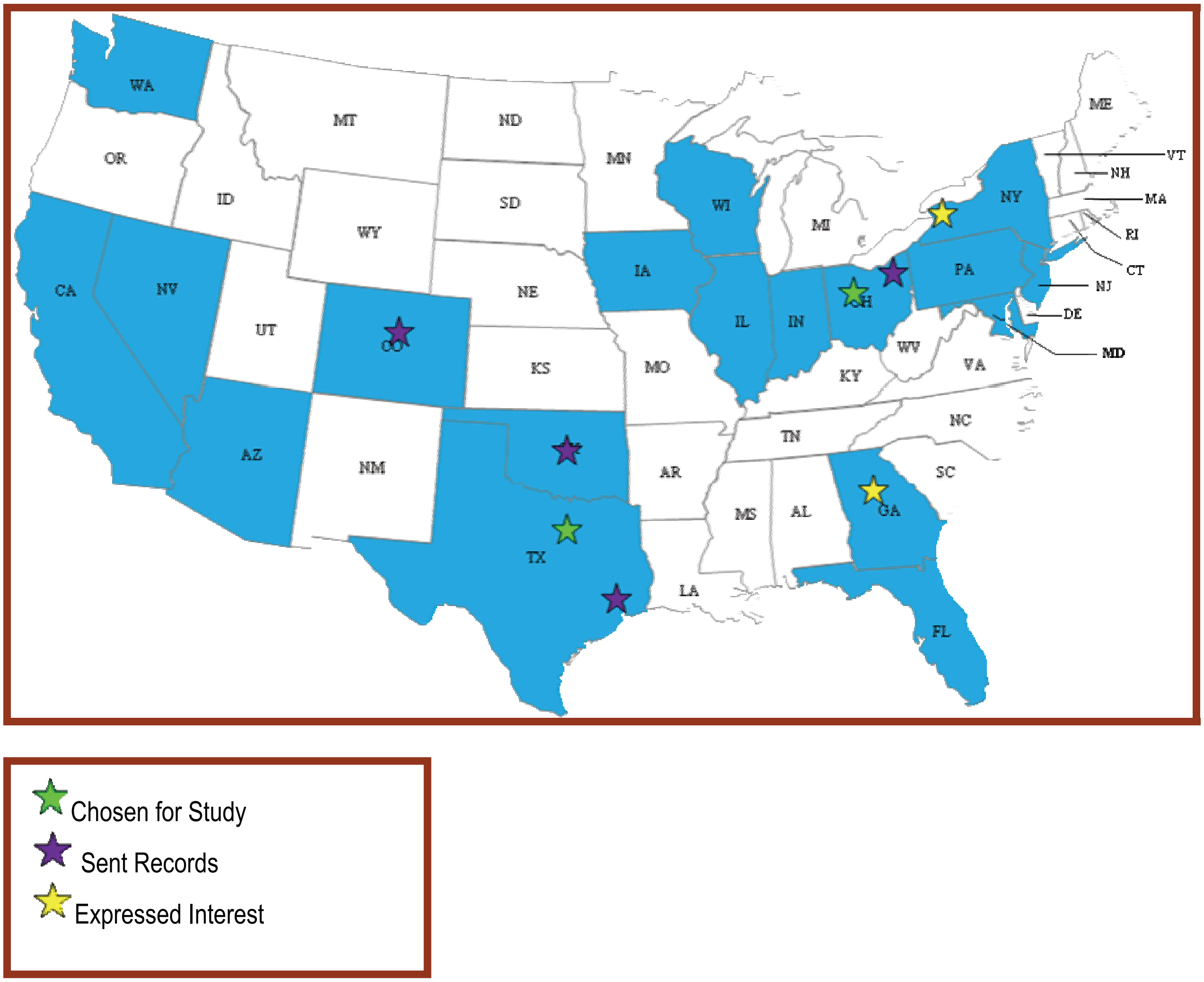
- Cleveland, OH
- Columbus, OH
- Denver, CO
- Oklahoma City, OK
- Dallas/Ft. Worth/Arlington, TX
- Houston, TX.

Because budget limits only allowed for a study of two MSAs, research was focused on the two MSAs from which the most Energy Star records were provided:

- Dallas/Ft. Worth/Arlington, TX
- Columbus, OH

The other MSAs either did not provide a robust enough sample, or the number of records available was simply lower than Dallas/Ft. Worth/Arlington and Columbus. Figure 5 shows the MSAs that showed interest, that actually sent records, and that were chosen for the study. Columbus records were provided by two HERS providers and Dallas/Ft. Worth/Arlington Records were provided by 3 HERS providers.

Figure 5



MSA Definitions:

Definitions for the Dallas/Ft. Worth/Arlington and Columbus MSAs are defined by the US Census Bureau and were derived from the Office of Management and Budget definitions for those MSAs. (OMB Bulletin No. 08-01 2007) This source lists the counties that make up the MSA. Because providers were often sending records from several MSAs, Newport used zip codes from these counties to eliminate records outside of the defined area. The counties included in the Dallas/Ft. Worth/Arlington, TX MSA are:

- Collin County
- Dallas County
- Delta County
- Denton County
- Ellis County
- Hunt County
- Johnson County
- Kaufman County
- Parker County

- Rockwall County
- Tarrant County
- Wise County

Figure 6 shows the MSA as defined by the Office of Management and Budget.

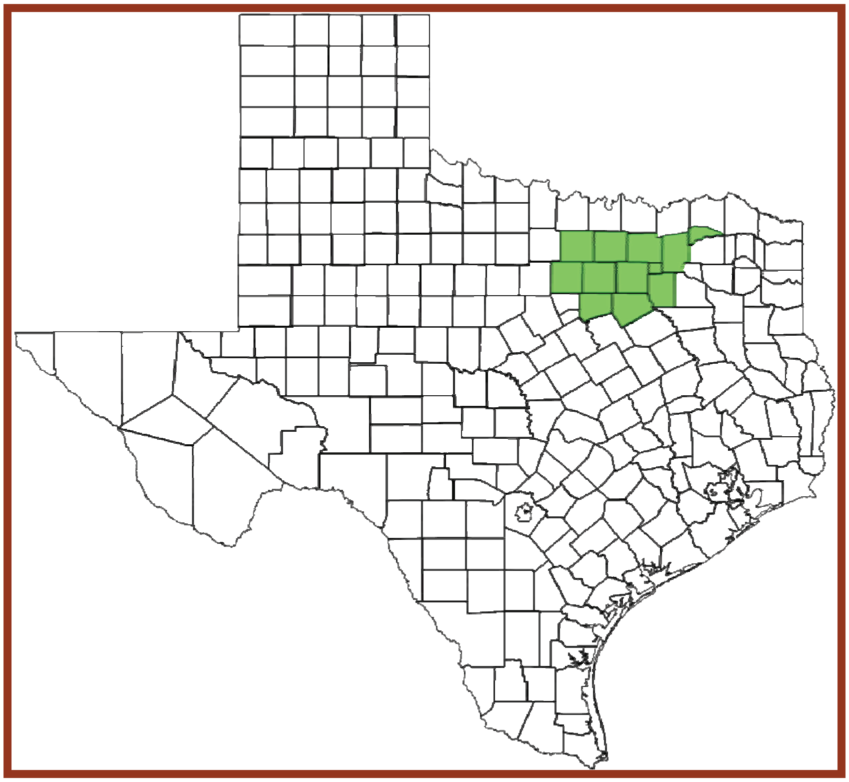


Figure 6

Note: Of these counties, no usable records were found in either the Energy Star group or the control group in Delta, Hunt, Johnson, Parker, and Wise after the data cleaning, matching, and filtering processes. The highest concentration of records in the Dallas/Ft. Worth/Arlington MSA was in Tarrant, Collin, Denton, and Dallas Counties.

The counties included in the Columbus, OH MSA were:

- Delaware County
- Fairfield County
- Franklin County
- Licking County
- Madison County
- Morrow County
- Pickaway County
- Union County

Figure 7 shows the MSA as defined by the Office of Management and Budget.

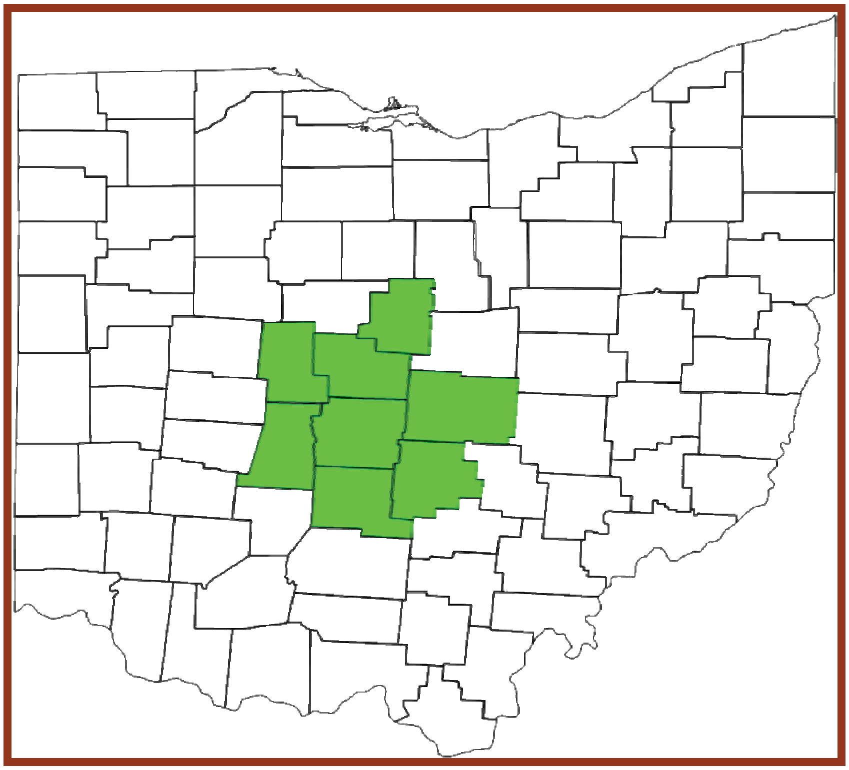


Figure 7

Note: Of these counties, no usable records were found in either the Energy Star group or the control group in Morrow County after the data cleaning, matching and filtering processes. The highest concentration of records was in Franklin and Delaware Counties.

Filtering Usable Energy Star Records:

In some cases records from providers were compiled into one spreadsheet. In others, multiple spreadsheets had to be combined and checked for duplicates. Some providers had already sorted Energy Star qualified homes from homes that received a HERS rating but did not qualify. When this had not been done, Newport sorted out all homes that did not meet Energy Star HERS requirements. Prior to 2006, an Energy Star home had to reach a HERS Score of 86 or above to qualify. The HERS baseline home was a HERS 80. In this system, higher HERS Scores were more efficient and lower HERS Scores less efficient. In 2006, RESNET changed the HERS Score to a HERS Index. On the HERS Index, Energy Star homes must meet a HERS 85 or lower. A HERS 100 was the baseline home, tied to the 2004 IECC. Under the HERS Index, lower numbers were more efficient and higher numbers less efficient. When filtering records from providers, any homes rated prior to 2006 achieving a less than a HERS Score of 86 were eliminated. From 2006 on, any homes achieving anything greater than a HERS Index of 85 were eliminated. Although some states chose not to change to the HERS Index in 2006, the study did not use data from any of those states. Any records that were incomplete in some way, such as missing vital address information; not showing HERS Index or HERS Score, etc. were filtered out. In the case of the Dallas/Ft. Worth/Arlington, TX MSA, homes in the City of Frisco, TX were eliminated from the sample. The City of Frisco requires that all homes be built to Energy Star standards. Therefore, the control group in this area would have been skewed. This initial filtering process produced 56,751 records in Texas and 15,990 in Ohio. These numbers were further cut down when they were sorted by zip code and determined to be outside of the required MSA counties.

Lessons Learned from Energy Star Records Collection Process:

1. Currently there is no national compilation of Energy Star records. Such a national database would greatly enhance the ability of researchers to collect data on Energy Star for a study that would include enough MSAs to be able to generalize the results. Without this centralization, researchers must spend greater time and resources pursuing individual HERS providers.
2. There is a wide variety of opinions from HERS providers about their data. Most considered the data their proprietary product. Some were interested in the project and happy to help. Others wanted approval from their builders before participating. Others wanted confidentiality or non-disclosure agreements signed. Some believed it would violate privacy laws for them to participate. The willingness to sign any confidentiality or non-disclosure agreements, and the endorsement of RESNET and Energy Star were vital to securing data from the HERS providers. In short, HERS providers saw no clear value in participation and were reluctant.
3. There is no incentive for HERS providers to participate in a study like this. Unless they were interested in the issue, or just liked being helpful, there was no motivation to provide data.
4. The format of the records varied greatly between HERS providers, as well as within a HERS provider's dataset. Without a national database, extensive formatting is necessary to use the Energy Star records, and to identify how many addresses they actually contained.
5. RESNET has indicated that it is developing a national registry of all HERS rated homes. Such a central database would make data collection much easier as it would remove the HERS provider from the position of having to decide whether or not to participate. RESNET could make the decision as an organization. One database would also ensure that records were in a common format. Access to the registry would also allow for greater flexibility in choosing MSAs to study.

Sales Transaction Records:

Several data sources for sales transaction records were considered including:

- Public records held by individual counties;
- Multiple Listing Service (MLS) data;
- Several different private data providers.

Several individual counties were examined during the study in order to determine whether or not it would be practical to collect transaction data directly from the counties. Many counties have transaction records publically available for free. However, they are readily available only on a record-by-record basis. Batches of records were available in some counties; however, collecting these batches would have involved purchasing them from the county, or going to the county record office and photocopying records. The labor intensiveness of this project, as well as the cost made this process prohibitive.

MLS as a data source was an attractive option because it is often available for purchase in most areas and would be easy to compile. However, MLS data will often not include the original sale of a property because real estate agents may or may not be used by the builder or developer during the original sale of a home. As this was an integral piece of data in the repeat sales index, MLS as a data source was not usable.

Several private data providers were examined that collect transaction data from county sources and/or MLS data sources and compile large quantities of data for use in research. Although all records would have to be rented from the provider, the work of collecting the data from individual counties would have been already performed. This last approach was selected. Like MLS, some providers did not have the original sale information and were eliminated. A nationally recognized data provider was selected that could provide transaction data, as well as the required logistical data about the home.

Selection of a private data provider was a long and careful process in which Newport sought advice from industry experts including advisory panel members; spoke with vendors about the project and about how their data matched with the needs of the project; and obtained quotes and data samples. A difficulty presented in the process is that finding a data provider is not just like choosing any vendor. Because researchers only rent data – obtain a license to use the data for a certain amount of time and/or a certain purpose – vendors are extremely protective of this data. The process involved not only discussions about the data provider's capabilities and products but also an in-depth discussion of the project. This was necessary to ensure the data provider that their product would not be compromised, nor would the use of their product lead to liability for the company because of the results of the study. Newport and the data provider went through multiple drafts of a licensing agreement to ensure that the data provider's product was protected and their liability was limited while still allowing Newport to maintain complete autonomy in the study and final report.

Data Filters:

The following data filters were requested from the data provider to ensure that only useful records were being paid for (see Table 3):

Table 3

Data Filter	Reasoning
Homes built 1995 or later *Note: the dataset was later restricted to homes built in 2001 or later and transaction pairs where the first sale occurred in 2002 or later due to a lack of significant Energy Star sample before that period.*	The energy efficient houses being used in the study were Energy Star homes. The repeat sales index needed a control group of homes that were built at the same time. Because the Energy Star for Homes program began in 1995, Newport used this year as its starting point for the control group.
Only homes that had been resold	A repeat sales index tracks the change in value of a home between sales events. Therefore at least one resale (two total sales) is necessary for a property to be used in the index.
Single Family Detached	This filter was put in place to establish similarity across the sample and also because most Energy Star homes would fall into this category.
Within the requested MSA as defined by county	The control group would necessarily be a part of the same MSA as the energy efficient homes to ensure a valid comparison. Also, because the study required transaction data for the energy efficient homes as well, this data would need to be matched to the energy efficient sample.
Include the entire dataset that meets these criteria	The study needed the entire dataset for each MSA because it would ensure that homes in the energy efficient sample would have transaction records and would not be missed as a result of only collecting a partial dataset. It also identified which homes among the energy efficient sample were single family homes and had been resold – making them usable in the study.

Sales Transaction Records Obtained:

The long process of obtaining sales transaction records ended with the private data provider leasing 811,581 sales transaction records to Newport, including 193,343 Columbus, and 618,238 in Dallas/Ft. Worth/Arlington. However, these numbers were individual sales transactions, not individual properties, and many were found to be built before 1995. The data matching process would later reduce this number dramatically.

Lessons Learned in the Data Collection Process:

1. Defining terms is extremely important in any data collection. In the case of this study, Newport considered a "record" to be one unique address including all of the sales of that property. This was important because Newport needed to know the number of addresses being used in the study for statistical analysis. The data provider considered "record" to mean a transaction record of an individual sale. This resulted in two problems. First it initially appeared as if the dataset contained many more addresses than it actually did. Next it led the data provider to focus on transactions that occurred 1995 or later rather than homes built 1995 or later. Newport was forced to reassess the data after receiving it and obtain a refund commensurate with the smaller dataset. Absolute precision in communications to and from the data provider is necessary to avoid collecting unusable or incomplete data.

2. Communication with the data provider's technical team in addition to their sales representative is vital, especially when using the data for complicated research. There were several instances where a technical question was asked about the data and the sales representative gave incorrect answers that were later corrected by the technical staff.

3. Be prepared to have a smaller set of usable data than what the total numbers originally represented by the dataset. Part of the nature of data collection and analysis is that records are often incomplete or incorrect. This can be a result of both errors in the original data source or errors in data collection and organization. Data are only a good as its source, and the source for data is county transaction records; any errors, typos, or omissions in those records will cause problems for whoever compiles this data. The steps Newport took to filter out incomplete, invalid, or questionable data in order to have as accurate a dataset as possible are detailed later in this paper.

4. One important note about data obtained from the Dallas/Ft. Worth/Arlington, TX MSA is that Texas is a "non-disclosure" state in which sale prices on property transactions are not listed in records. In these states, the data provider uses a proprietary formula based on the mortgage amount for the home and multiplied by a factor based on the loan type. Although the data provider was not willing to release the exact nature of the formula due to its proprietary nature, they were willing to provide this general overview of how it worked. It is unlikely that this would cause problems in the repeat sales index because both the Energy Star and control group homes would be using this same formula in Texas.

Data Matching Process:

The next step in the data process was to clean the data and match the Energy Star records with sales transaction records. This action – performed by Abt Associates (Abt) - accomplished several objectives. First, it standardized all of the data. Next, it combined Energy Star addresses with sales transaction data, and it also eliminated the Energy Star addresses not found in the transaction data. This ensured that Energy Star houses that did not meet the MSA, single-family detached, or resold at least once criteria were not included in the analysis. The data cleaning and matching process contained four major steps.

1. **Compiling Address Data** – This step included taking data files of sale transactions and Energy Star homes and compiling the key address data. Data were cleaned in preparation for geocoding.

2. **Geocoding of Address Data** – Address data were analyzed using geocoding software, which included address standardization and location output data. Through geocoding, Abt could determine which addresses were located in the analysis MSAs.

3. **Address Matching** – Using the standardized address outputs, particularly the extended ZIP Code data, addresses in the lists of Energy Star homes were matched to the file of sale transactions.

4. **Creation of Final Analysis Files** – After address matching, once the sale transactions were flagged as Energy Star homes, the sale transactions data were processed to create property-level and sale transactions-level files of home with at least two sale transactions.

Compiling Address Data

The first set of input data included a file of properties from a private data provider. The original data file contained 811,581 records (618,238 records in Texas and 193,343 records in Ohio). In addition to property address, year built, and sale transactions information – including sale date, recording date, and sale price – the data file had owner contact information and selected property characteristics, including lot size, living area, number of bedrooms, number of baths, number of stories, and zoning. While starting the cleaning process, Abt discovered a problem with the sales transaction data. It was not a quality problem but a classification problem. What Newport had expected to be 811,581 addresses with transaction records included were actually 811,581 transaction records, representing far fewer addresses; and, many records listed the year built of the home as being prior to 1995. After eliminating all homes built prior to 1995, there remained 347,626 sales transaction records in the file, including 70,226 in Ohio and 277,400 in Texas. All records were submitted for address standardization and geocoding.

The second set of input data included files of Energy Star addresses. At a minimum, the files contained address information, and some files contained other data, including builder names and HERS Scores and HERS Index values. In all, there were 15,990 address records compiled from the Ohio HERS providers and 56,751 address records compiled from the Texas HERS providers. All records were submitted for address standardization and geocoding.

Once the address data were compiled, the addresses were cleaned where possible. This included creating separate fields for the street address, city, state, and ZIP Code, checking the spelling of city names, and checking city names against the ZIP Codes. Both the sale transactions address file and the energy rated homes address file included what appeared to be a duplicate address data. All records were submitted for address standardization and geocoding to confirm location in the analysis metropolitan areas. Unique address records were determined at a later stage of data processing.

Geocoding of Address Data

Data files were entered into software for address standardization and geocoding. All addresses were geocoded using Group 1 CODE-1 plus v0.3.2. All address data were standardized using US Postal Service data files vintage 4/1/2008.

The key data for the address matching task were the address output fields, including the standardized street address, city name, and ZIP Code fields. The ZIP Code fields include the five-digit ZIP Code, the four-digit ZIP Plus 4 Code, and the two-digit Delivery Point Code. The four-digit ZIP Plus 4 Code is an extension of the ZIP Code identifying a smaller segment of the five-digit ZIP Code postal delivery area. The two-digit delivery point code is another extension of the ZIP Code, identifying a specific unique delivery address. Together, the five-digit ZIP Code, four-digit ZIP Plus 4 Code, and the two-digit Delivery Point Code make up an eleven-digit extended ZIP Code field. If the geocoding software was able to output a standardized address with complete street address and extended ZIP Code fields, the record was flagged as geocoded. Addresses in the sale transactions data file were also checked for delivery point validation. If the address could not be validated, the address standardization output was deleted, and the address was considered not geocoded.

The geocoding software was able to standardize and output geocoding data for over 90% of the address records submitted. With the implied state location of Ohio or Texas, an address could be geocoded with a recognizable input street address and either city name or ZIP Code. Records that did not geocode were often missing both the ZIP Code and city name. A summary of the geocoding rates by data file is shown in Table 4.

Table 4

Geocoding Rates for Input Address Data Files			
Address Data File	Number of Records Submitted	Number of Records Geocoded	Geocoding Rate
Sale Transactions Data	347,626	342,119	98.4 percent
Ohio Energy Star Homes	15,990	15,103	94.5 percent
Texas Energy Star Homes	56,751	52,764	93.0 percent

Once the address data were geocoded, records were selected based on location in the analysis MSAs. Abt then determined unique addresses based on the extended ZIP code fields. The majority of geocoded sale transactions address records were geocoded to either the Columbus, OH MSA or the Dallas/Ft. Worth/Arlington, TX MSA, meaning that most of the data received from data providers were in the correct MSAs. Table 5 summarizes the number of geocoded records in the input address files, the number of records geocoded to the analysis MSAs, and the number of unique address records in the final files for address matching. A small portion of the sale transactions records that geocoded outside of the analysis MSAs were located primarily in non-metropolitan area locations. Of the 340,293 sale transactions records geocoded to the analysis MSAs, there were 185,999 unique addresses after combining multiple sales transactions of a single address.

Table 5

Summary of Geocoded Address Data Available for Address Matching – Sale Transactions			
	Number of Records Geocoded	Number of Records in the Analysis MSAs	Number of Unique Addresses
All Sale Transactions Data	342,119	340,293	184,659
Columbus, OH MSA		65,961	33,247
Dallas/Ft. Worth/Arlington, TX MSA		274,332	151,412

Notes: There were 1,846 address records (0.5%) that did not geocode to either the Columbus, OH MSA or the Dallas/Ft. Worth/Arlington, TX MSA.

Table 6 summarizes the number of geocoded records in the input address files of Energy Star homes, the number of records geocoded to the analysis MSAs, and the number of unique address records in the final files for address matching. Nearly 95% of geocoded address records provided by the Ohio HERS providers were located in Columbus, OH MSA. About 80% of the address records provided by the Texas HERS Providers were located in the Dallas/Ft. Worth/Arlington, TX MSA. In total, there were 10,492 unique addresses in the Columbus, OH MSA and 33,262 unique addresses in the Dallas/Ft. Worth/Arlington, TX MSA. Duplicate addresses were known to be in and across the initial data sources.

Table 6

Summary of Geocoded Address Data Available for Address Matching – Energy Star Homes			
	Number of Records Geocoded	Number of Records in the Analysis MSAs	Number of Unique Addresses
Columbus, OH MSA	15,103	14,288	10,492
Dallas/Ft. Worth/Arlington, TX MSA	52,764	42,317	33,262

Notes: There were 815 address records (5.4%) that did not geocode to the Columbus, OH MSA and 10,447 address records (19.8%) that did not geocode to the Dallas/Ft. Worth/Arlington, TX MSA.

Address Matching

Once the data files of unique addresses were created, the next step was to complete the address matching. The task involved flagging the addresses in the sale transactions file that were also in the file of Energy Star homes.

The address matching methodology we undertook uses the extended ZIP Code data. This methodology is most appropriate when the address data are for single family properties. The extended ZIP Code data include the nine-digit ZIP Code and the two-digit delivery point code. The delivery point code, which is based on the house number, identifies unique delivery locations within a nine-digit ZIP Code. Multi-unit buildings have multiple delivery locations, and the US Postal Service will not recognize a delivery point code for the main address of a multi-unit building. Instead, each unit or delivery point within the multi-unit building will be assigned a delivery point code based on the unit number. Because the properties in this address matching task are single unit or single family properties, each with its own delivery point, address matching using the extended ZIP Code was an appropriate methodology choice.

The same geocoding system was used to standardize the addresses, thus the basis for the geocoding output was consistent throughout all of the address data files. Because the eleven-digit extended ZIP Code created a unique identifier for an address, this matching methodology would provide definitive matches. We would also be more likely to find matches using the extended ZIP Codes than matching on the standardized address data fields. In some areas, the US Postal Service recognizes more than one city name for a given street address. Compared to matching only on the address data fields, matching on the extended ZIP Code would account for any differences in city name and allow for more numbers of accurate matches.

Matching was completed using the extended ZIP Code fields and the files of unique sale transactions addresses and unique energy rated homes addresses. If an address was in the sale transactions file and the Energy Star homes file, the address was flagged as a match in the sale transactions file. If an address was in the sale transactions file but not in the energy rated homes file, the address was flagged as not a match in the sale transactions file.

Table 7 summarizes the initial results of address matching. Of the 33,247 unique addresses of sale transactions in the Columbus, OH MSA, 2,249 were energy rated homes. Of the 151,412 unique addresses of sale transactions in the Dallas/Ft. Worth/Arlington, TX MSA, 7,255 were Energy Star homes. The address matches were confirmed by reviewing the standardized street address data. In the sale transactions file, the remaining addresses – 30,998 in the Columbus, OH MSA and 144,157 in the Dallas/Ft. Worth/Arlington, TX MSA – that did not match to energy rated homes were flagged as non-matched records. These non-matched records would be used as the non-energy efficient control group.

Table 7

Initial Results of Address Matching
Energy Star Homes Addresses and Sale Transactions Addresses

	Number of Unique Energy Star Home Addresses	**Number of Unique Sale Transactions Addresses**	**Number of Address Matches**
Columbus, OH MSA	10,492	33,247	2,249
Dallas/Ft. Worth/Arlington, TX MSA	33,262	151,412	7,255

While the main focus of the address matching was to find sale transactions for energy rated homes, it is important to note that of the unique addresses of Energy Star homes in the Columbus, OH MSA and Dallas/Ft. Worth/Arlington, TX MSA files, only about one-fifth were matched to sale transactions records. It is likely that the other Energy Star records that were not matched with transaction records did not meet one of the necessary criteria such as having been resold or being a single-family detached house.

Creation of Final Analysis Files

Once the address matching was complete, the next steps were to pull all the sale transactions and create the final files. The final files included sale transactions data matched to Energy Star homes and sale transactions data not matched to Energy Star homes. To create files of properties with multiple sale transactions, the final files included only the addresses with more than one sale transaction record in the geocoded data file. Addresses with only one sale transaction were dropped from data processing into the final data files. Any addresses with only one sale transaction would not meet the required criteria of having been resold.

Using the data on matched addresses, all sale transactions for the matched addresses were compiled. A series of data processing and checks were done on the sale transactions data, including:

- Removing sale transactions that were complete duplicates of other sale transaction records.
- Keeping only the property addresses with more than one sale transaction.
- Deleting property addresses with inconsistent property characteristics across sale transactions, since changes in property characteristics may affect sale price.

Table 8 summarizes the property-level final data processing for the matched sale transactions records. The summaries are shown by MSA. For the Columbus, OH MSA, geocoding found 33,247 unique sale transaction addresses. After completing address matching using the extended ZIP Code, we found 2,249 properties were energy rated homes. Nearly 500 of these properties were deleted because there were data on only one sale transaction. After checking across property sales records for inconsistencies, another 15 properties were deleted. The final data file of sale transactions of Energy Star homes in the Columbus, OH MSA included 1,742 properties. The same data processing steps were used in compiling the Dallas/Ft. Worth/Arlington, TX MSA data. The final data file of sale transactions of energy rated homes in the Dallas/Ft. Worth/Arlington, TX MSA included 3,040 properties.

Table 8

Summary Results of Address Matching
Property-Level Sale Transactions of Energy Star Homes

	Number of Unique Sale Transactions Addresses	Number of Energy Star Address Matches	Number of Properties with Only One Sale Transaction	Number of Properties with Inconsistent Property Data	Number of Energy Star Properties in Final File
Columbus, OH MSA	33,247	2,249	492	15	**1,742**
Dallas/Ft. Worth/Arlington, TX MSA	151,412	7,255	4,189	26	**3,040**

The same data processing was done to sale transaction records that did not match to the file of Energy Star homes. Table 9 summarizes the property-level final data processing for the non-matched sale transactions records. Of the 33,247 unique sale transaction addresses geocoded to the Columbus, OH MSA, 30,998 were flagged as not matching to the Energy Star homes address file. There were 7,642 properties deleted because there were data on only one sale transaction. After checking across property sales records for inconsistencies, another 214 properties were deleted. The final data file of sale transactions not matching to the file of Energy Star homes in the Columbus, OH MSA resulted in a control group of 23,142 properties. The same data processing steps were used in compiling the Dallas/Ft. Worth/Arlington, TX MSA data file of non-matching properties, resulting in a control group of 94,502 properties.

Table 9

Summary Results of Non-Matching Addresses
Property-Level Sale Transactions Not Matched to Energy Star Addresses

	Number of Unique Sale Transactions Addresses	Number of Addresses Not Matched to File of Energy Star Homes	Number of Properties with Only One Sale Transaction	Number of Properties with Inconsistent Property Data	Number of Control Group Properties in Final File
Columbus, OH MSA	33,247	30,998	7,642	214	23,142
Dallas/Ft. Worth/Arlington, TX MSA	151,412	144,157	48,776	876	94,502

The final files are comprised of data from sale transactions and from the geocoding software. For each MSA, Abt created two files for both Energy Star records and control group records. Each group included a property level file, including all sales for each property. Each group also included sales transaction files with separate records for each transaction. After the cleaning, geocoding and matching process, Abt returned the files to Newport for use in the repeat sales index.

Repeat Sales Index Process and Results

Even after the previously described processes and editing of the property and sales records, there needed to be further work cleaning the data to eliminate transaction pairs with prices or price changes that were inappropriate or inaccurate and that would have distorted the results. Examples of inappropriate transactions include land sales and sales between related parties. In some cases such inappropriate transactions were detectable because the purchaser was a builder, the "year built" reported for the current structure was later than the sale date, or the buyer and seller had the same name. In other cases, records were excluded because the reported values or changes were implausibly extreme. The exclusion rules based on extreme values were somewhat arbitrary, but the likelihood that the excluded observations were legitimate and appropriate appeared to be less than the likelihood that they were faulty. Other repeat sales measures, such as OFHEO and Case-Shiller, use similar, but not identical, exclusion criteria.

The data for the calculation were restricted to homes built in 2001 or later and transaction pairs where the first sale occurred in 2002 or later. This is a shorter period than would have been preferred, but the available data included only 3 Energy Star homes in Texas purchased before 2002, and from both locations there was only a handful of records for Energy Star homes built before 2001.

Table 10 shows the exclusion rules and the effect of successive application of those rules on the number of observations in each of the 4 classes.

Table 10

Exclusion Rules

	Surviving Pairs				
	Columbus		Dallas/Ft. Worth/Arlington		All Groups
	Energy	Control	Energy	Control	
Initial Properties	***1,742***	***23,142***	***3,040***	***94,502***	***122,426***
Initial Sales Pairs	***2,058***	***29,968***	***3,228***	***117,805***	***153,059***
Pairs Built 2001 or later, Sold 2002 or later	**1,577**	**5,475**	**3,216**	**32,761**	**43,029**
Builder / Corporation	727	2,084	1,979	17,013	21,803
Buyer and seller same person	710	2,013	1,950	16,777	21,450
Record Date before sale or >180 days after	699	1,970	1,938	16,598	21,205
1st and 2nd sale of a pair within same quarter	684	1,912	1,925	16,353	20,874
Invalid date	684	1,912	1,924	16,348	20,868
Year built after purchase year	681	1,858	1,920	15,994	20,453
Price < $95,000 or > $1,000,000	671	1,793	1,840	14,964	19,268
Annualized appreciation > 50% or < -50%	665	1,769	1,812	14,604	18,850
Total appreciation > 100% or < -50%	665	1,767	1,788	14,416	18,636
Final Sample	**665**	**1,767**	**1,788**	**14,416**	**18,636**
	Eliminated by Test (marginal exclusions)				
	Columbus		**Dallas/Ft. Worth/Arlington**		**All Groups**
	Energy	**Control**	**Energy**	**Control**	
Builder / Corporation	850	3,391	1,237	15,748	21,226
Buyer and seller same person	17	71	29	236	353
Record Date before sale or >180 days after	11	43	12	179	245
1st and 2nd sale of a pair within same quarter	15	58	13	245	331
Invalid date	0	0	1	5	6
Year built after purchase year	3	54	4	354	415
Price < $95,000 or > $1,000,000	10	65	80	1,030	1,185
Annualized appreciation > 50% or < -50%	6	24	28	360	418
Total appreciation > 100% or < -50%	0	2	24	188	214
	Failing Test				
	Columbus		**Dallas/Ft. Worth/Arlington**		**All Groups**
	Energy	**Control**	**Energy**	**Control**	
Builder / Corporation	850	3,391	1,237	15,748	21,226
Buyer and seller same person	21	100	36	402	559
Record Date before sale or >180 days after	43	242	20	415	720
1st and 2nd sale in a pair within same quarter	170	734	137	1,654	2,695
Invalid date	1	7	1	11	20
Year built after purchase year	222	913	103	3,144	4,382
Price < $95,000 or > $1,000,000	308	1,413	383	4,657	6,761
Annualized appreciation > 50% or < -50%	477	2,334	526	6,625	9,962
Total appreciation > 100% or < -50%	373	1,926	437	4,379	7,115

Most exclusions occurred because they included purchases by builders, generally of land prior to construction. In the data file, there was a field for the full name of the buyer, and separate fields for first and last names of the primary and secondary (generally spouse) buyers. Where there were no first and last names, it was found that the full name was generally a business name. In some cases, the business was not a builder, but instead a real estate company, relocation service, or institution. These were excluded as well because of concerns (supported by other anomalies) that they were not arms-length transactions at market value.

The other tests, including unusually high or low values or rates of appreciation, had more limited effects on the size of the samples. Although there were quite a few records that did not pass those tests, they were already eliminated by the exclusion of purchases by builders and other disqualifications.

More than one pair of sales could be obtained for a property that sold more than twice during the 2002-2007 period, but there were few cases of properties involved in multiple pairs, as is evident from the relatively small difference between the number of properties and the number of pairs in Table 11. Overall, there were 18,636 sales pairs used, from 17,937 properties.

Table 11

Sales Pairs and Properties

	Columbus Energy	Control	Dallas/Ft. Worth/Arlington Energy	Control	All Samples
Total Properties	643	1,701	1,748	13,845	17,937
Total Sales Pairs	665	1,767	1,788	14,416	18,636
Properties with					
1 Pair	621	1,637	1,709	13,284	17,251
2 Pairs	22	62	38	551	673
3 Pairs	0	2	1	10	13

Tables 12-16 show the distribution of prices, appreciation rates, and other characteristics for the Energy Star and control groups in the two locations before and after the exclusion rules were applied.

Characteristics of Homes in Sales-Pair Samples

1st Sale in a Transaction Pair

	AFTER EXCLUSIONS				BEFORE EXCLUSIONS			
	Columbus, OH		Dallas/Ft. Worth/Arlington, TX		Columbus, OH		Dallas/Ft Worth/Arlington, TX	
	Energy Star	Control	Energy Star	Control	Energy Star	Control	Energy Star	Control
Sales Pairs	665	1,767	1,788	14,416	1,577	5,475	3,216	32,761
Mean	**246,329**	**257,205**	**200,851**	**226,335**	**468,815**	**303,826**	**3,643,712**	**1,130,279**
Std Deviation	92,668	124,949	77,980	120,490	668,115	326,263	38,100,000	9,758,201
Minimum	116,975	97,000	95,028	95,025	25,000	1,000	4,123	2,584
Percentiles:								
1	131,400	119,000	100,413	101,146	34,000	30,000	73,150	39,401
5	143,800	133,900	118,477	114,805	76,875	49,500	109,588	93,312
10	154,700	144,800	127,619	124,205	118,076	67,500	123,639	113,715
25	178,500	170,700	147,693	146,034	169,900	128,000	146,451	144,576
50 (Median)	**221,900**	**216,600**	**179,676**	**188,276**	**240,000**	**205,300**	**186,746**	**203,490**
75	291,300	307,200	229,636	266,000	349,500	339,000	263,307	329,840
90	356,800	420,000	307,163	376,001	1,185,500	635,000	452,200	579,547
95	453,300	515,000	357,304	466,830	2,193,000	1,000,000	20,000,000	950,950
99	542,400	730,000	462,724	707,493	2,674,000	1,882,400	53,200,000	33,300,000
Maximum	740,000	936,000	811,699	992,047	4,375,000	2,597,600	1,200,000,000	1,200,000,000
Skewness	1.44	1.80	1.86	2.13	3.01	2.91	28.83	61.59
Kurtosis	5.53	7.01	8.65	9.03	12.77	13.31	899.26	6,957.63

2nd Sale in a Transaction Pair

	AFTER EXCLUSIONS				BEFORE EXCLUSIONS			
	Columbus, OH		Dallas/Ft. Worth/Arlington, TX		Columbus, OH		Dallas/Ft. Worth/Arlington, TX	
	Energy Star	Control	Energy Star	Control	Energy Star	Control	Energy Star	Control
Sales Pairs	665	1,767	1,788	14,416	1,577	5,475	3,216	32,761
Mean	**243,983**	**258,649**	**213,093**	**240,792**	**278,670**	**317,838**	**398,311**	**373,663**
Std Deviation	106,442	140,212	89,488	130,474	157,215	223,634	2,682,476	1,507,507
Minimum	95,000	95,000	95,060	95,000	1,000	1,100	13,300	421
Percentiles:								
1	101,313	102,501	102,332	100,930	84,500	54,000	74,812	69,160
5	118,000	116,000	118,769	116,933	122,000	114,002	102,521	104,405
10	129,000	127,500	128,744	128,866	136,900	132,100	119,700	121,828
25	161,000	154,500	154,183	156,674	172,400	173,000	148,302	155,610
50 (Median)	**225,000**	**220,000**	**191,121**	**199,500**	**246,900**	**257,000**	**186,865**	**212,800**
75	303,000	319,500	240,647	280,896	330,000	397,000	244,720	341,145
90	377,500	444,000	335,052	403,056	471,600	575,000	347,462	571,900
95	474,000	560,000	392,350	508,725	537,000	691,000	426,930	832,646
99	545,000	738,500	510,720	754,110	733,000	1,312,000	851,200	1,862,000
Maximum	750,000	980,000	948,290	997,500	1,710,000	3,336,600	53,200,000	53,200,000
Skewness	1.12	1.57	2.05	2.11	3.03	2.87	15.56	22.70
Kurtosis	4.37	5.86	10.48	8.66	21.22	18.55	259.37	562.55

Table 12

Characteristics of Homes in Sales-Pair Samples

Price Relative

	AFTER EXCLUSIONS				BEFORE EXCLUSIONS			
	Columbus, OH		Dallas/Ft Worth/Arlington, TX		Columbus, OH		Dallas/Ft Worth/Arlington, TX	
	Energy Star	Control	Energy Star	Control	Energy Star	Control	Energy Star	Control
Sales Pairs	665	1,767	1,788	14,416	1,577	5,475	3,216	32,761
Mean	**0.97**	**0.99**	**1.08**	**1.09**	**1.23**	**2.08**	**1.52**	**1.51**
Std Deviation	0.12	0.14	0.23	0.24	1.31	2.89	8.40	6.38
Minimum	0.62	0.53	0.52	0.50	0.00	0.00	0.00	0.00
Percentiles:								
1	0.68	0.68	0.59	0.58	0.13	0.13	0.00	0.01
5	0.75	0.74	0.71	0.71	0.16	0.25	0.01	0.39
10	0.79	0.79	0.79	0.78	0.21	0.47	0.49	0.65
25	0.89	0.90	0.91	0.92	0.82	0.89	0.84	0.89
50 (Median)	**1.00**	**1.01**	**1.07**	**1.07**	**1.00**	**1.04**	**1.04**	**1.08**
75	1.05	1.07	1.23	1.25	1.09	1.97	1.24	1.30
90	1.11	1.14	1.38	1.40	1.75	5.75	1.43	1.60
95	1.14	1.18	1.46	1.48	4.47	6.53	1.61	2.91
99	1.22	1.34	1.65	1.74	7.11	9.23	8.40	8.55
Maximum	1.28	1.96	1.93	2.00	10.57	113.60	274.29	400.24
Skewness	-0.50	0.30	0.27	0.32	3.53	12.95	25.87	39.66
Kurtosis	2.75	5.50	2.98	3.14	16.86	425.23	740.63	1,937.96

Annual Relative

	AFTER EXCLUSIONS				BEFORE EXCLUSIONS			
	Columbus, OH		Dallas/Ft Worth/Arlington, TX		Columbus, OH		Dallas/Ft Worth/Arlington, TX	
	Energy Star	Control	Energy Star	Control	Energy Star	Control	Energy Star	Control
Sales Pairs	665	1,767	1,788	14,416	1,538	5,308	3,195	32,470
Mean	**0.99**	**1.00**	**1.03**	**1.03**	**2.06E+10**	**1.32E+34**	**2.98E+17**	**2.65E+32**
Std Deviation	0.06	0.07	0.12	0.12	7.94E+11	8.14E+35	1.69E+19	4.77E+34
Minimum	0.82	0.71	0.54	0.50	0.00	0.00	0.00	0.00
Percentiles:								
1	0.87	0.84	0.72	0.68	0.00	0.00	0.00	0.00
5	0.91	0.90	0.83	0.83	0.09	0.13	0.00	0.23
10	0.93	0.93	0.88	0.89	0.26	0.38	0.51	0.72
25	0.96	0.96	0.96	0.97	0.92	0.94	0.91	0.94
50 (Median)	**1.00**	**1.00**	**1.03**	**1.03**	**1.00**	**1.02**	**1.02**	**1.04**
75	1.03	1.03	1.09	1.09	1.05	2.14	1.11	1.16
90	1.05	1.06	1.16	1.15	2.59	11.36	1.29	1.71
95	1.07	1.10	1.22	1.22	10.99	34.25	1.73	2.76
99	1.12	1.24	1.38	1.38	99.12	3834.20	14.00	40.11
Maximum	1.49	1.47	1.49	1.50	3.11E+13	5.87E+37	9.53E+20	8.59E+36
Skewness	1.05	1.21	0.11	-0.02	39.16	70.86	56.50	180.19
Kurtosis	13.16	10.43	4.57	5.13	1,535.18	5,100.81	3,193.00	32,468.00

Table 13

Characteristics of Homes in Sales-Pair Samples								
	Year Built							
	AFTER EXCLUSIONS				**BEFORE EXCLUSIONS**			
	Columbus, OH		**Dallas/Ft. Worth/Arlington, TX**		**Columbus, OH**		**Dallas/Ft Worth/Arlington, TX**	
	Energy Star	**Control**	**Energy Star**	**Control**	**Energy Star**	**Control**	**Energy Star**	**Control**
Sales Pairs	665	1,767	1,788	14,416	1,577	5,475	3,216	32,761
Mean	**2002.58**	**2002.36**	**2003.42**	**2002.58**	**2002.99**	**2002.94**	**2003.63**	**2003.15**
Std Deviation	1.01	1.06	1.10	1.20	1.15	1.23	1.13	1.49
Minimum	2001	2001	2001	2001	2001	2001	2001	2001
Percentiles:								
1	2001	2001	2001	2001	2001	2001	2001	2001
5	2001	2001	2002	2001	2001	2001	2002	2001
10	2001	2001	2002	2001	2002	2001	2002	2001
25	2002	2002	2003	2002	2002	2002	2003	2002
50 (Median)	**2002**	**2002**	**2004**	**2002**	**2003**	**2003**	**2004**	**2003**
75	2003	2003	2004	2003	2004	2004	2004	2004
90	2004	2004	2005	2004	2005	2005	2005	2005
95	2004	2004	2005	2005	2005	2005	2005	2006
99	2005	2005	2005	2006	2005	2005	2006	2006
Maximum	2005	2005	2006	2006	2006	2006	2006	2006
Skewness	0.47	0.58	-0.19	0.55	0.12	0.18	-0.28	0.31
Kurtosis	2.72	2.69	2.32	2.65	2.12	2.16	2.32	2.09

Table 14

Characteristics of Homes in Sales-Pair Samples

Baths

	AFTER EXCLUSIONS				BEFORE EXCLUSIONS			
	Columbus, OH		Dallas/Ft Worth/Arlington, TX		Columbus, OH		Dallas/Ft Worth/Arlington, TX	
	Energy Star	Control	Energy Star	Control	Energy Star	Control	Energy Star	Control
Sales Pairs	665	1,767	1,788	14,416	1,577	5,475	3,216	32,761
Mean	**2.99**	**3.08**	**2.52**	**2.58**	**3.11**	**3.28**	**2.48**	**2.75**
Std Deviation	0.53	0.63	0.73	0.83	0.95	0.94	0.75	1.08
Minimum	2	1	1	1	2	1	1	1
Percentiles:								
1	2	2	2	2	2	2	2	2
5	2	2	2	2	2	2	2	2
10	2	2	2	2	2	3	2	2
25	3	3	2	2	3	3	2	2
50 (Median)	**3**	**3**	**2**	**2**	**3**	**3**	**2**	**2**
75	3	3	3	3	3	4	3	3
90	4	4	4	4	4	4	4	4
95	4	4	4	4	4	5	4	5
99	5	5	4	5	5	6	4	7
Maximum	6	6	8	32	33	26	8	32
Skewness	0.65	1.09	1.47	4.11	19.74	5.19	1.41	2.33
Kurtosis	6.48	6.16	6.41	111.94	614.82	95.00	5.62	23.67

Bedrooms

	AFTER EXCLUSIONS				BEFORE EXCLUSIONS			
	Columbus, OH		Dallas/Ft Worth/Arlington, TX		Columbus, OH		Dallas/Ft Worth/Arlington, TX	
	Energy Star	Control	Energy Star	Control	Energy Star	Control	Energy Star	Control
Sales Pairs	665	1,767	1,788	14,416	1,577	5,475	3,216	32,761
Mean	**3**	**4**	**4**	**4**	**4**	**4**	**4**	**4**
Std Deviation	1	1	1	1	1	1	1	1
Minimum	1	2	2	2	1	1	2	1
Percentiles:								
1	2	2	3	3	2	2	2	3
5	3	3	3	3	3	3	3	3
10	3	3	3	3	3	3	3	3
25	3	3	3	3	3	3	3	3
50 (Median)	**3**	**4**	**4**	**4**	**4**	**4**	**4**	**4**
75	4	4	4	4	4	4	4	4
90	4	4	5	5	4	4	5	5
95	4	4	5	5	4	4	5	5
99	4	5	5	5	4	5	5	5
Maximum	5	6	11	16	5	7	11	16
Skewness	-0.28	-0.07	1.01	1.33	-0.34	-0.14	0.85	1.01
Kurtosis	2.41	2.58	9.01	13.22	2.51	3.68	6.38	7.29

Table 15

Characteristics of Homes in Sales-Pair Samples

Lot Size (acres)

	AFTER EXCLUSIONS				BEFORE EXCLUSIONS			
	Columbus, OH		Dallas/Ft Worth/Arlington, TX		Columbus, OH		Dallas/Ft Worth/Arlington, TX	
	Energy Star	Control	Energy Star	Control	Energy Star	Control	Energy Star	Control
Sales Pairs	441	1,250	1,249	11,171	974	3,703	2,345	26,458
Mean	**0.21**	**0.30**	**0.20**	**2.54**	**0.23**	**0.43**	**0.19**	**2.57**
Std Deviation	0.10	0.42	0.48	164.15	0.34	0.90	0.36	154.41
Minimum	0.07	0.02	0.08	0.00	0.07	0.02	0.05	0.00
Percentiles:								
1	0.07	0.11	0.09	0.09	0.07	0.10	0.08	0.09
5	0.11	0.12	0.12	0.12	0.11	0.13	0.12	0.12
10	0.13	0.14	0.13	0.13	0.12	0.14	0.12	0.13
25	0.14	0.17	0.14	0.14	0.15	0.17	0.14	0.15
50 (Median)	**0.18**	**0.22**	**0.17**	**0.17**	**0.19**	**0.25**	**0.16**	**0.18**
75	0.24	0.29	0.20	0.23	0.25	0.36	0.20	0.26
90	0.32	0.43	0.25	0.34	0.37	0.72	0.25	0.50
95	0.38	0.60	0.29	0.62	0.45	1.28	0.30	1.00
99	0.50	1.91	0.50	1.50	0.79	3.85	0.65	2.00
Maximum	1.03	5.57	16.00	15572.00	10.05	23.47	16.00	15572.00
Skewness	2.76	7.80	30.10	85.03	25.31	13.98	38.15	83.83
Kurtosis	17.95	78.11	977.82	7,639.72	731.02	293.99	1,649.02	7,564.17

Living Area (sq ft)

	AFTER EXCLUSIONS				BEFORE EXCLUSIONS			
	Columbus, OH		Dallas/Ft Worth/Arlington, TX		Columbus, OH		Dallas/Ft Worth/Arlington, TX	
	Energy Star	Control	Energy Star	Control	Energy Star	Control	Energy Star	Control
Sales Pairs	665	1,767	1,788	14,416	1,577	5,475	3,216	32,761
Mean	**2,208**	**2,315**	**2,499**	**2,536**	**2,388**	**2,598**	**2,468**	**2,731**
Std Deviation	682	697	733	862	762	891	777	1,146
Minimum	1,072	976	1,064	900	1,072	576	991	798
Percentiles:								
1	1,167	1,200	1,271	1,243	1,173	1,200	1,157	1,150
5	1,282	1,350	1,524	1,445	1,300	1,432	1,408	1,343
10	1,393	1,514	1,639	1,578	1,490	1,582	1,584	1,519
25	1,749	1,792	1,916	1,853	1,810	1,910	1,868	1,860
50 (Median)	**2,036**	**2,182**	**2,377**	**2,352**	**2,208**	**2,478**	**2,344**	**2,495**
75	2,657	2,768	3,004	3,085	2,938	3,164	2,993	3,391
90	3,182	3,305	3,574	3,799	3,510	3,741	3,596	4,200
95	3,544	3,578	3,840	4,153	3,674	4,116	3,921	4,810
99	3,860	4,225	4,317	4,900	4,302	5,350	4,344	6,325
Maximum	6,383	5,458	4,927	7,142	6,383	7,856	7,068	14,036
Skewness	0.92	0.72	0.57	0.79	0.58	0.95	0.63	1.38
Kurtosis	4.53	3.27	2.67	3.17	2.86	4.60	3.10	7.02

Table 16

In the Columbus samples (after exclusions), the prices of first and second sales in a transaction pair for the Energy Star homes were slightly higher than the medians for the control group, with the median first sale price of $221,900 versus $216,600. In both locations, however, average prices were higher for the control groups, reflecting the fact that the variation in prices was greater for the control groups – the high end was higher and the low end lower. That is reflected in the standard deviations. The control groups were also more heterogeneous in terms of characteristics such as the number of bathrooms and living area. In general, however, the structural characteristics and market values of the Energy Star and control groups were similar.

Some of the characteristic values, even after exclusions, are not reasonable. For example, the maximum numbers for lot size are as high as 15,572 acres. That's probably a case where square feet rather than acres were recorded. For some properties, data for lot size were not available. In any case, those characteristics don't enter into the repeat sales calculation, and most of the data on characteristics are reasonable, with similar values for the Energy Star and control groups.

With inappropriate and suspicious records removed, regressions were run to estimate quarterly index values for each of the four classes as shown in Table 17.

Index Regressions by Class

	Columbus, OH MSA								Dallas/Ft. Worth/Arlington, TX MSA							
	OH Energy				OH Control				TX Energy				TX Control			
	obs=665		R2=0.213		obs=1767		R2=0.107		obs=1788		R2=0.113		obs=14416		R2=0.120	
	Coeff	Std Err	t-stat	Index	Coeff	Std Err	t-stat	Index	Coeff	Std Err	t-stat	Index	Coeff	Std Err	t-stat	Index
2002Q2	-0.012	0.024	-0.50	98.81	0.026	0.015	1.73	102.61	0.057	0.041	1.38	105.83	0.026	0.009	2.95	102.61
2002Q3	-0.010	0.022	-0.46	98.98	0.019	0.015	1.29	101.95	0.077	0.039	1.95	107.97	0.027	0.009	3.04	102.76
2002Q4	-0.029	0.023	-1.26	97.12	0.020	0.015	1.31	102.02	0.061	0.041	1.48	106.28	0.025	0.009	2.83	102.55
2003Q1	-0.038	0.025	-1.52	96.31	0.033	0.017	1.96	103.33	0.030	0.039	0.77	103.06	0.018	0.009	1.94	101.83
2003Q2	-0.055	0.027	-2.02	94.63	0.003	0.016	0.22	100.34	0.049	0.036	1.36	105.01	0.035	0.009	3.70	103.53
2003Q3	-0.059	0.024	-2.41	94.31	0.005	0.016	0.33	100.53	0.068	0.037	1.87	107.08	0.029	0.009	3.19	102.93
2003Q4	-0.045	0.023	-1.96	95.60	-0.008	0.016	-0.48	99.24	0.076	0.036	2.14	107.93	0.042	0.009	4.63	104.32
2004Q1	-0.039	0.027	-1.47	96.13	-0.011	0.017	-0.63	98.94	0.041	0.037	1.10	104.18	0.039	0.009	4.15	104.01
2004Q2	0.003	0.024	0.11	100.27	0.003	0.016	0.18	100.28	0.111	0.035	3.19	111.73	0.077	0.009	8.52	108.03
2004Q3	-0.016	0.025	-0.62	98.45	0.027	0.015	1.73	102.69	0.107	0.034	3.12	111.25	0.093	0.009	10.76	109.77
2004Q4	-0.017	0.024	-0.70	98.30	-0.012	0.016	-0.73	98.81	0.104	0.034	3.03	110.93	0.105	0.009	11.31	111.08
2005Q1	-0.030	0.027	-1.09	97.06	0.014	0.017	0.85	101.42	0.103	0.036	2.90	110.86	0.108	0.009	11.52	111.41
2005Q2	-0.004	0.024	-0.18	99.57	0.028	0.015	1.88	102.86	0.158	0.034	4.62	117.12	0.130	0.009	15.09	113.83
2005Q3	-0.017	0.023	-0.76	98.28	0.006	0.016	0.41	100.64	0.172	0.034	4.99	118.79	0.130	0.008	15.53	113.86
2005Q4	-0.046	0.024	-1.93	95.48	0.000	0.016	-0.02	99.97	0.138	0.035	3.91	114.84	0.137	0.009	15.29	114.68
2006Q1	-0.075	0.026	-2.84	92.79	-0.046	0.018	-2.59	95.54	0.173	0.036	4.84	118.90	0.139	0.009	15.39	114.97
2006Q2	-0.057	0.023	-2.43	94.50	-0.010	0.015	-0.68	98.98	0.144	0.034	4.20	115.51	0.135	0.008	16.24	114.48
2006Q3	-0.064	0.024	-2.65	93.76	0.001	0.015	0.09	100.13	0.166	0.034	4.81	118.00	0.142	0.008	17.20	115.25
2006Q4	-0.104	0.025	-4.19	90.14	-0.029	0.016	-1.89	97.11	0.166	0.035	4.73	118.06	0.136	0.009	15.40	114.52
2007Q1	-0.118	0.023	-5.12	88.87	-0.079	0.016	-4.79	92.44	0.167	0.035	4.80	118.12	0.123	0.009	13.92	113.11
2007Q2	-0.101	0.022	-4.56	90.38	-0.044	0.014	-3.15	95.66	0.168	0.033	5.04	118.35	0.123	0.008	14.94	113.14
2007Q3	-0.101	0.024	-4.26	90.35	-0.026	0.014	-1.77	97.47	0.142	0.034	4.22	115.22	0.096	0.008	11.88	110.11
2007Q4	-0.149	0.027	-5.45	86.12	-0.116	0.018	-6.43	89.04	0.045	0.037	1.22	104.55	0.041	0.010	3.98	104.17

Table 17

The indices were each set to 100 for the first quarter of 2002. The indices for the Energy Star and control samples in each of the two locations are shown in Figures 8 and 9.

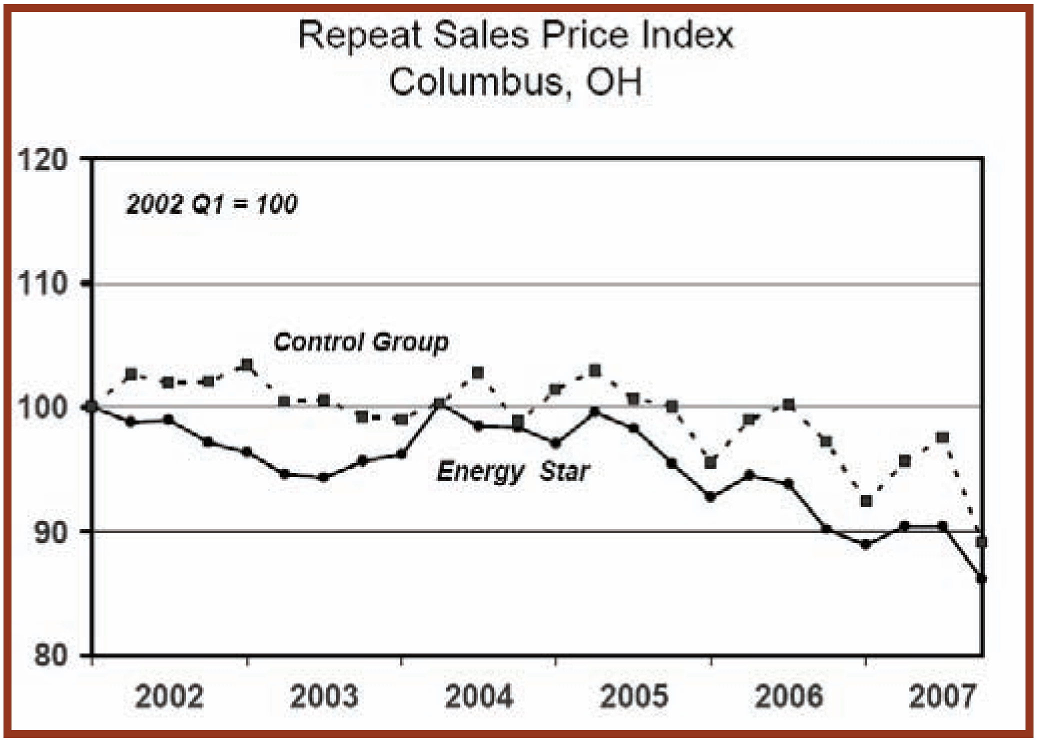


Figure 8

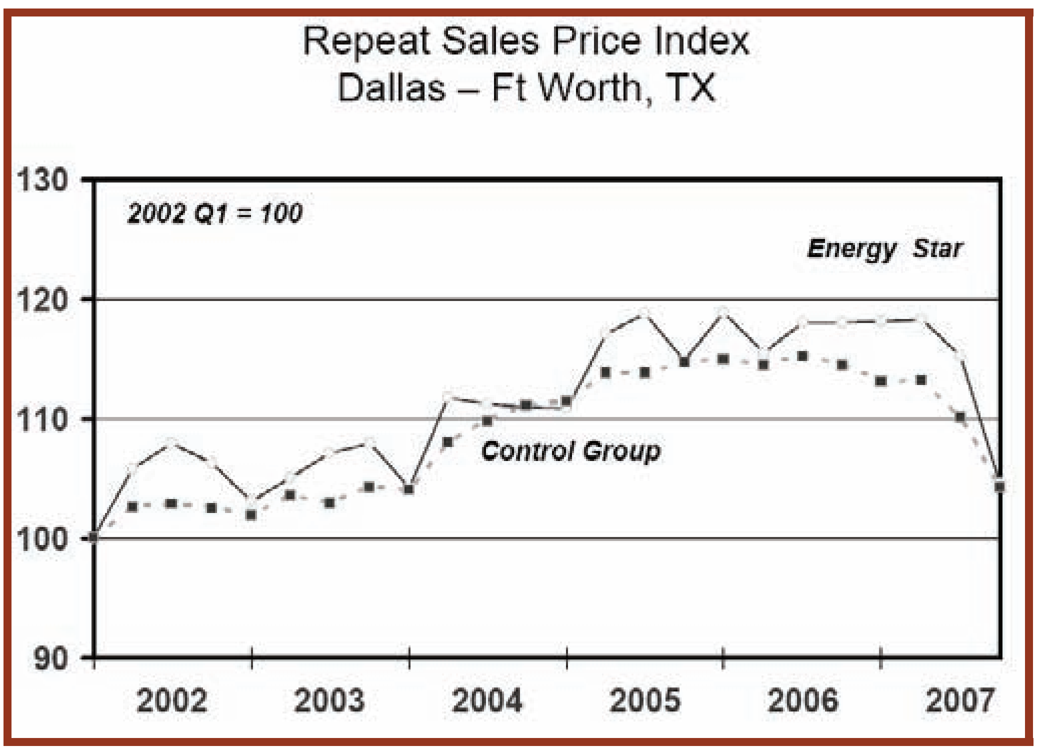


Figure 9

For the Columbus MSA, the index for the control group was generally higher than for the Energy Star group. For the Dallas/Ft. Worth/Arlington MSA, however, the Energy Star index was generally higher than the control group index. The apparent relative strength viewed in this way is sensitive to the starting point. If the indices for Columbus used 2003Q1 as a base, for example, rather than 2002Q1, the appreciation through the fourth quarter of 2007 would appear to be higher for Energy Star homes, as shown in Figure 10.

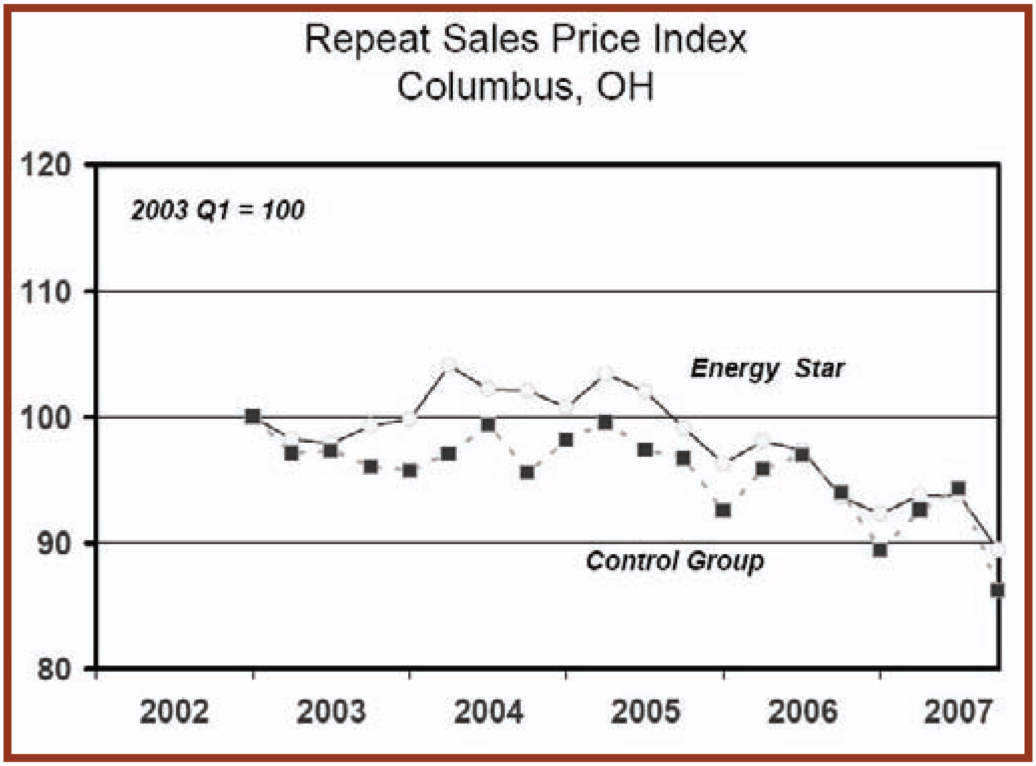


Figure 10

The alternative repeat sales calculation of quarterly appreciation rates, rather than indices benchmarked to a specified base period, provides a more useful test of whether there are consistent differences in appreciation (although even then it matters what time range is used to determine whether pairs are included), as shown in Table 18.

Period Relative Regressions by Class

	Columbus, OH MSA								Dallas/Ft. Worth/Arlington, TX MSA							
	OH Energy				OH Control				TX Energy				TX Control			
	obs=665			R2=0.213	obs=1767			R2=0.107	obs=1788			R2=0.113	obs=14416			R2=0.120
	Coeff	Std Err	t-stat	% chg	Coeff	Std Err	t-stat	% chg	Coeff	Std Err	t-stat	% chg	Coeff	Std Err	t-stat	% chg
2002Q2	-0.012	0.024	-0.50	-1.19%	0.026	0.015	1.73	2.61%	0.057	0.041	1.38	5.83%	0.026	0.009	2.95	2.61%
2002Q3	0.002	0.021	0.08	0.17%	-0.006	0.015	-0.43	-0.64%	0.020	0.036	0.56	2.02%	0.001	0.009	0.17	0.15%
2002Q4	-0.019	0.020	-0.96	-1.88%	0.001	0.015	0.05	0.07%	-0.016	0.036	-0.44	-1.56%	-0.002	0.009	-0.22	-0.20%
2003Q1	-0.008	0.022	-0.37	-0.83%	0.013	0.017	0.75	1.28%	-0.031	0.036	-0.85	-3.03%	-0.007	0.009	-0.75	-0.70%
2003Q2	-0.018	0.027	-0.64	-1.75%	-0.029	0.018	-1.68	-2.89%	0.019	0.031	0.61	1.90%	0.017	0.010	1.66	1.66%
2003Q3	-0.003	0.027	-0.12	-0.33%	0.002	0.017	0.11	0.19%	0.020	0.028	0.70	1.97%	-0.006	0.010	-0.60	-0.58%
2003Q4	0.014	0.022	0.62	1.37%	-0.013	0.017	-0.76	-1.28%	0.008	0.027	0.29	0.79%	0.013	0.009	1.42	1.35%
2004Q1	0.005	0.025	0.22	0.55%	-0.003	0.018	-0.17	-0.30%	-0.035	0.027	-1.32	-3.48%	-0.003	0.010	-0.30	-0.30%
2004Q2	0.042	0.026	1.62	4.31%	0.013	0.018	0.75	1.36%	0.070	0.026	2.71	7.25%	0.038	0.010	3.83	3.87%
2004Q3	-0.018	0.025	-0.73	-1.81%	0.024	0.017	1.41	2.40%	-0.004	0.022	-0.20	-0.43%	0.016	0.009	1.74	1.61%
2004Q4	-0.002	0.025	-0.06	-0.15%	-0.039	0.017	-2.22	-3.78%	-0.003	0.022	-0.13	-0.29%	0.012	0.009	1.27	1.19%
2005Q1	-0.013	0.026	-0.48	-1.26%	0.026	0.018	1.41	2.64%	-0.001	0.024	-0.02	-0.06%	0.003	0.010	0.30	0.30%
2005Q2	0.025	0.027	0.93	2.58%	0.014	0.017	0.81	1.42%	0.055	0.023	2.42	5.64%	0.021	0.009	2.28	2.17%
2005Q3	-0.013	0.023	-0.56	-1.29%	-0.022	0.016	-1.35	-2.16%	0.014	0.022	0.65	1.43%	0.000	0.008	0.03	0.03%
2005Q4	-0.029	0.022	-1.30	-2.85%	-0.007	0.016	-0.40	-0.66%	-0.034	0.024	-1.42	-3.33%	0.007	0.009	0.82	0.72%
2006Q1	-0.029	0.026	-1.11	-2.82%	-0.045	0.019	-2.42	-4.43%	0.035	0.025	1.39	3.53%	0.002	0.009	0.26	0.25%
2006Q2	0.018	0.025	0.72	1.84%	0.035	0.018	1.93	3.60%	-0.029	0.024	-1.22	-2.85%	-0.004	0.009	-0.48	-0.42%
2006Q3	-0.008	0.023	-0.34	-0.78%	0.012	0.015	0.75	1.17%	0.021	0.022	0.97	2.16%	0.007	0.008	0.84	0.67%
2006Q4	-0.039	0.025	-1.58	-3.87%	-0.031	0.016	-1.90	-3.02%	0.001	0.023	0.02	0.05%	-0.006	0.009	-0.75	-0.63%
2007Q1	-0.014	0.024	-0.60	-1.40%	-0.049	0.018	-2.81	-4.80%	0.000	0.023	0.02	0.05%	-0.012	0.009	-1.38	-1.23%
2007Q2	0.017	0.021	0.81	1.70%	0.034	0.016	2.14	3.47%	0.002	0.021	0.09	0.19%	0.000	0.008	0.03	0.03%
2007Q3	0.000	0.021	-0.02	-0.03%	0.019	0.014	1.36	1.89%	-0.027	0.018	-1.46	-2.64%	-0.027	0.008	-3.53	-2.67%
2007Q4	-0.048	0.026	-1.83	-4.68%	-0.090	0.018	-5.12	-8.64%	-0.097	0.023	-4.22	-9.26%	-0.055	0.010	-5.65	-5.39%

Table 18

As noted above, the calculations of quarterly appreciation and of price indices produce the same estimates of change. If the calculated quarterly appreciation rates are successively applied to the base period value, the result is the index. Conversely, the relative changes from quarter to quarter in the estimated index are identical to the estimates of quarterly appreciation.

Quarterly appreciation rates for the Energy Star and control groups in the 2 locations are shown in Figures 11 and 12. For Columbus, in 11 of the 23 quarters covered, the estimated appreciation for the Energy Star homes was greater than for the control group. For Dallas/Ft. Worth/Arlington, the estimated Energy Star appreciation was higher in 13 quarters. Thus, in neither case was there a consistent advantage found for either group.

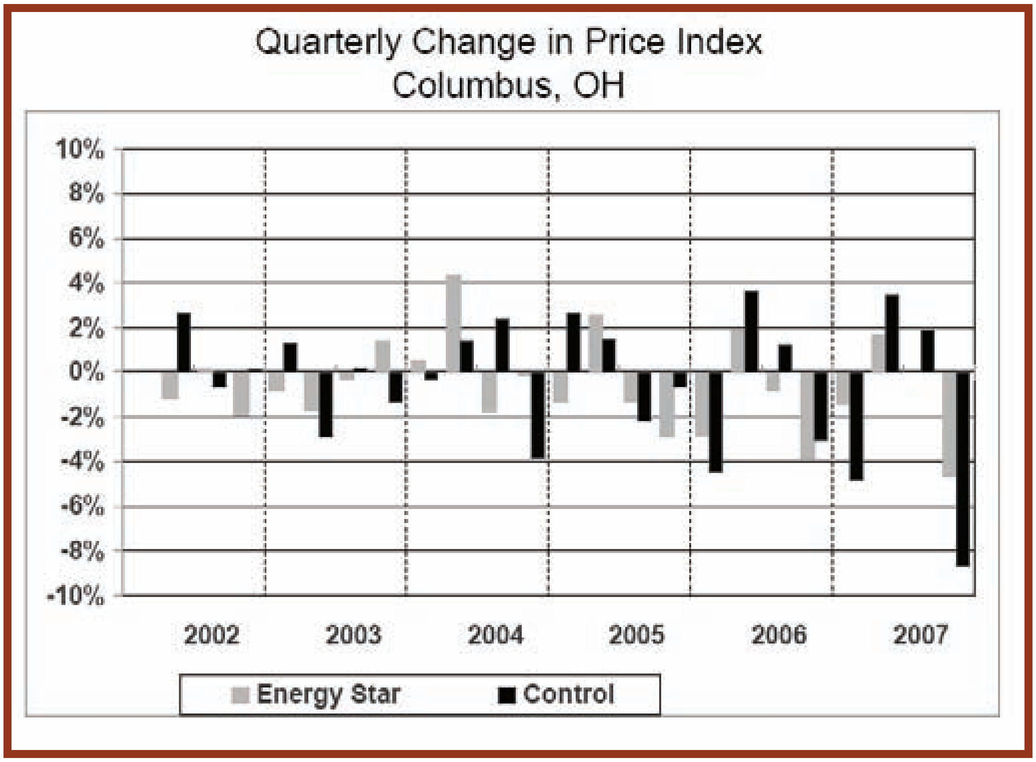


Figure 11

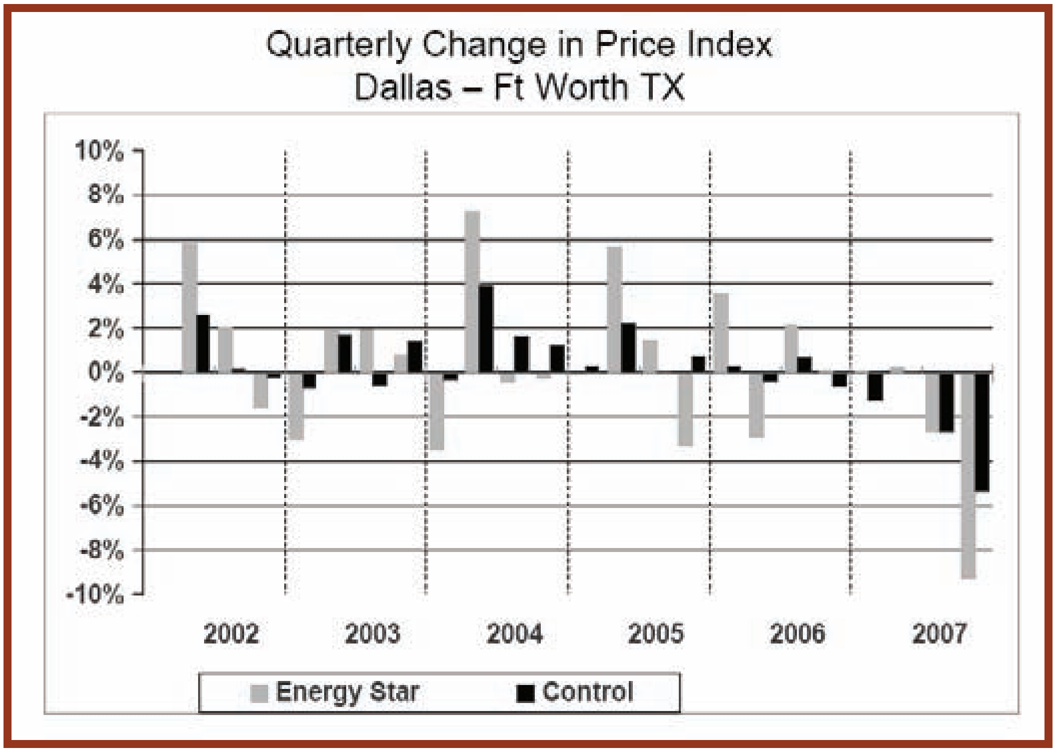


Figure 12

The R-squared values for all these regressions are unimpressive. In part that is because there was little movement in overall price levels in either of the two MSAs over the time range studied, so area-wide change for each group didn't explain much of whatever changes occurred for individual houses. There may also have been heterogeneity within groups, producing dissimilar experience, and perhaps a more uniform sample was needed. Or there could have been undetected errors in the data.

To more formally test whether appreciation for Energy Star homes was greater or less than for the control group to a statistically significant extent, the Energy Star and control groups were combined into a single regression, for each MSA, with interaction/slope dummies. The results are shown in Table 19. The data values for the slope dummy variables are the same (0 or 1) as standard dummies for Energy Star homes, but are always equal to zero for the control group homes. For the control group, the coefficients of the standard dummies are identical to those calculated in the regression based only on those homes (e.g., 0.0135 for 2004Q2 in Columbus), but for the Energy Star homes the coefficients on the slope dummies plus the coefficients for the standard dummies are equal to coefficients for the separate regression(e.g., 0.0135+0.0287=0.0422 for 2004Q2). Under this arrangement, the question of whether differences in the appreciation rates are statistically significant is measured by whether the slope dummies are significantly different from zero. For individual quarters, the standard test of significance at the 5% level would be based on whether the ratios of the slope dummies to their standard errors (the t-statistics) are greater than 1.96. With 23 quarterly slope dummies for each location, a 5 percent chance could be expected to translate into 1 or 2 "significant" differences, but none of the t-statistics exceed 1.62 in absolute value.

Period Relative Regressions with Slope Dummies for Energy Star

	Columbus MSA					Dallas-Ft Worth-Arlington MSA				
	Number of obs = 2432	F(46, 2386) = 7.93		Prob > F = 0.0000		Number of obs = 16.204		F(46. 16158) = 47.59		Prob > F = 0.0000
	R-squared = 0.1326	Adi R-squared = 0.1158		Root MSE = 0.1365		R-squared = 0.1193		Adi R-squared = 0.1168		Root MSE = 0.2201
	Coef.	**S.E.**	**t**	**P>|t|**	***Pct Chg***	**Coef.**	**S.E.**	**t**	**P>|t|**	***Pct Chg***
	Standard Dummies					**Standard Dummies**				
2002Q2	0.0257	0.0143	1.79	0.073	*-2.61%*	0.0257	0.0087	2.96	0.003	*-2.61%*
2002Q3	-0.0064	0.0143	-0.45	0.654	*0.64%*	0.0015	0.0087	0.17	0.864	*-0.15%*
2002Q4	0.0007	0.0147	0.05	0.962	*-0.07%*	-0.0020	0.0089	-0.22	0.822	*0.20%*
2003Q1	0.0128	0.0164	0.78	0.437	*-1.28%*	-0.0070	0.0094	-0.75	0.454	*0.70%*
2003Q2	-0.0293	0.0169	-1.73	0.083	*2.89%*	0.0165	0.0099	1.67	0.096	*-1.66%*
2003Q3	0.0019	0.0163	0.11	0.909	*-0.19%*	-0.0058	0.0097	-0.60	0.546	*0.58%*
2003Q4	-0.0129	0.0165	-0.78	0.435	*1.28%*	0.0134	0.0094	1.43	0.154	*-1.35%*
2004Q1	-0.0030	0.0173	-0.17	0.861	*0.30%*	-0.0030	0.0098	-0.30	0.764	*0.30%*
2004Q2	0.0135	0.0174	0.78	0.438	*-1.36%*	0.0379	0.0099	3.85	0.000	*-3.87%*
2004Q3	0.0238	0.0163	1.46	0.144	*-2.40%*	0.0159	0.0091	1.74	0.081	*-1.61%*
2004Q4	-0.0385	0.0168	-2.30	0.022	*3.78%*	0.0118	0.0093	1.27	0.203	*-1.19%*
2005Q1	0.0261	0.0179	1.46	0.145	*-2.64%*	0.0030	0.0099	0.30	0.762	*-0.30%*
2005Q2	0.0141	0.0168	0.84	0.404	*-1.42%*	0.0215	0.0094	2.29	0.022	*-2.17%*
2005Q3	-0.0218	0.0157	-1.39	0.164	*2.16%*	0.0003	0.0084	0.03	0.975	*-0.03%*
2005Q4	-0.0066	0.0158	-0.42	0.676	*0.66%*	0.0072	0.0088	0.82	0.412	*-0.72%*
2006Q1	-0.0454	0.0181	-2.50	0.012	*4.43%*	0.0025	0.0093	0.26	0.793	*-0.25%*
2006Q2	0.0354	0.0178	1.99	0.046	*-3.60%*	-0.0043	0.0088	-0.49	0.627	*0.42%*
2006Q3	0.0116	0.0149	0.78	0.436	*-1.17%*	0.0067	0.0080	0.84	0.401	*-0.67%*
2006Q4	-0.0307	0.0156	-1.97	0.049	*3.02%*	-0.0064	0.0085	-0.75	0.452	*0.63%*
2007Q1	-0.0492	0.0169	-2.91	0.004	*4.80%*	-0.0124	0.0090	-1.38	0.167	*1.23%*
2007Q2	0.0341	0.0155	2.21	0.027	*-3.47%*	0.0003	0.0085	0.03	0.974	*-0.03%*
2007Q3	0.0187	0.0134	1.40	0.161	*-1.89%*	-0.0271	0.0076	-3.54	0.000	*2.67%*
2007Q4	-0.0904	0.0171	-5.29	0.000	*8.64%*	-0.0554	0.0098	-5.68	0.000	*5.39%*
	Slope Dummies - Energy Star					**Slope Dummies - Energy Star**				
2002Q2	-0.0377	0.0302	-1.25	0.212	*3.70%*	0.0309	0.0433	0.71	0.475	*-3.14%*
2002Q3	0.0082	0.0270	0.30	0.762	*-0.82%*	0.0185	0.0380	0.49	0.627	*-1.87%*
2002Q4	-0.0197	0.0263	-0.75	0.455	*1.95%*	-0.0138	0.0384	-0.36	0.720	*1.37%*
2003Q1	-0.0211	0.0298	-0.71	0.478	*2.09%*	-0.0237	0.0386	-0.62	0.538	*2.35%*
2003Q2	0.0117	0.0348	0.34	0.737	*-1.18%*	0.0023	0.0334	0.07	0.945	*-0.23%*
2003Q3	-0.0052	0.0344	-0.15	0.881	*0.52%*	0.0254	0.0303	0.84	0.403	*-2.57%*
2003Q4	0.0265	0.0295	0.90	0.369	*-2.68%*	-0.0056	0.0293	-0.19	0.848	*0.56%*
2004Q1	0.0085	0.0323	0.26	0.793	*-0.85%*	-0.0324	0.0294	-1.10	0.270	*3.19%*
2004Q2	0.0287	0.0338	0.85	0.395	*-2.91%*	0.0321	0.0285	1.13	0.260	*-3.26%*
2004Q3	-0.0421	0.0321	-1.31	0.190	*4.12%*	-0.0203	0.0247	-0.82	0.412	*2.01%*
2004Q4	0.0370	0.0321	1.15	0.249	*-3.77%*	-0.0147	0.0248	-0.59	0.553	*1.46%*
2005Q1	-0.0388	0.0344	-1.13	0.259	*3.80%*	-0.0036	0.0265	-0.14	0.892	*0.36%*
2005Q2	0.0114	0.0347	0.33	0.743	*-1.15%*	0.0334	0.0253	1.32	0.187	*-3.40%*
2005Q3	0.0088	0.0301	0.29	0.770	*-0.88%*	0.0139	0.0242	0.58	0.564	*-1.40%*
2005Q4	-0.0223	0.0293	-0.76	0.447	*2.20%*	-0.0411	0.0262	-1.57	0.116	*4.02%*
2006Q1	0.0168	0.0339	0.49	0.621	*-1.69%*	0.0323	0.0275	1.17	0.240	*-3.28%*
2006Q2	-0.0171	0.0334	-0.51	0.608	*1.70%*	-0.0247	0.0260	-0.95	0.343	*2.44%*
2006Q3	-0.0194	0.0297	-0.65	0.513	*1.92%*	0.0146	0.0242	0.60	0.545	*-1.47%*
2006Q4	-0.0088	0.0318	-0.28	0.783	*0.87%*	0.0069	0.0251	0.27	0.784	*-0.69%*
2007Q1	0.0351	0.0312	1.12	0.261	*-3.58%*	0.0129	0.0257	0.50	0.615	*-1.30%*
2007Q2	-0.0173	0.0277	-0.63	0.531	*1.72%*	0.0017	0.0230	0.07	0.942	*-0.17%*
2007Q3	-0.0191	0.0271	-0.70	0.482	*1.89%*	0.0003	0.0205	0.02	0.987	*-0.03%*
2007Q4	0.0424	0.0338	1.26	0.209	*-4.34%*	-0.0417	0.0257	-1.62	0.105	*4.09%*
	F-test for Slope Dummies (all=0)	**F(23, 2386) = 1.05**		**Prob > F = 0.3936**		**F-test for Slope Dummies (all=0)**		**F(23, 16158) = 0.93**		**Prob > F = 0.5531**

Table 19

A better test, however, of whether the appreciation rates are significantly different is to consider whether the set of slope dummy coefficients could all be zero. F-test measures of that hypothesis indicate that there is no statistically significant difference between the Energy Star and control groups' appreciation rates in either location. For the slope dummies to be significantly different from zero at the 5% level, the F statistic would have to be greater than about 1.53. For Columbus, we have F(23/2386)=1.05, while for Dallas/Ft. Worth/Arlington, the result is F(23/16158)=0.93. With only two MSAs, the fact that there was no significant difference between the Energy Star and control groups (and that the small differences in the two areas pointed in opposite directions) provides weak support for the hypothesis that there is no difference in the appreciation rates. To confirm that there is no difference; data from a larger number of MSAs should be analyzed.

Comparison with Broader Indices

For Columbus, the repeat sales indices calculated in our analysis for both the Energy Star and control groups indicate weaker appreciation than for all homes in the metropolitan area. For Dallas/Ft. Worth/Arlington, price changes for the sample groups were similar to those from broader indices until the second half of 2006, after which price indices for the sample homes did worse compared to indices such as those from OFHEO, Case-Shiller, and Zillow. Figures 13 and 14 show comparisons with the OFHEO indices.

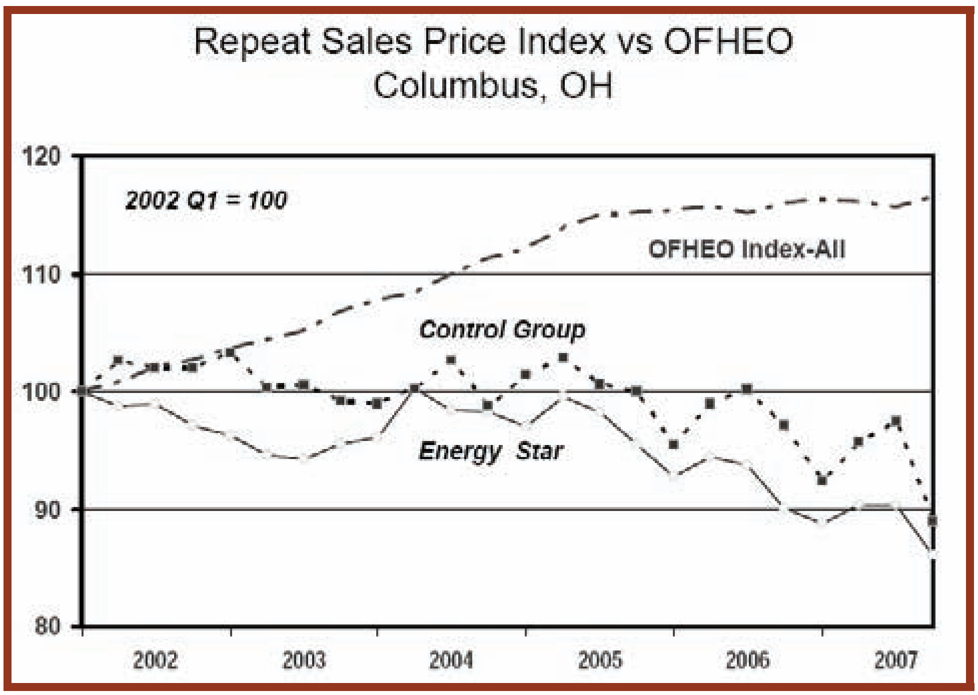


Figure 13

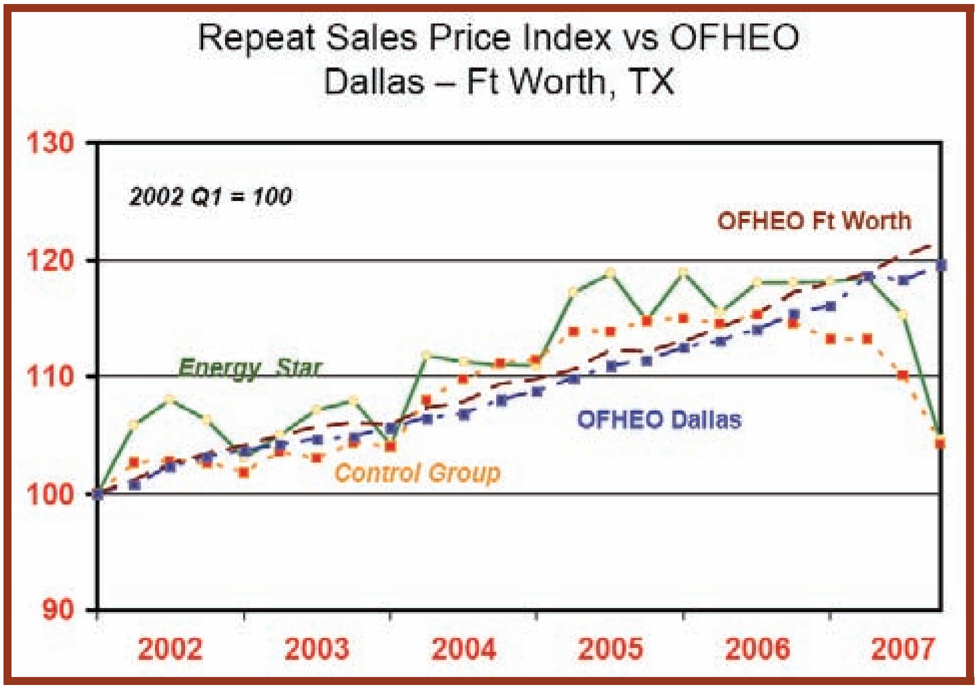


Figure 14

The weaker prices in our samples may reflect the fact that newer homes tend to be located in fringe areas. Several recent studies have shown that housing values at the fringe of metropolitan areas have been weaker than values in close-in neighborhoods. (OFHEO News Release 2007; Cortright 2008; Stiff 2008) It may also be true that values for homes built since 2000 have been weaker than values for older homes, even after considering distance from the metropolitan core. Although there are no reports (other than this one) showing price trends for recently built homes, data for 2007 from the Census Bureau's Housing Vacancy Survey showed the homeowner vacancy rate for homes built in 2000 or later as 8.1%, compared to an overall rate of 2.7%.

The apparent discrepancy between appreciation rates for newer homes and older homes underscores the importance of comparing appreciation of energy efficient homes to homes that are otherwise comparable, in terms of vintage, location, and structural characteristics, rather than to more general measures of price trends

Survey

The main focus of the study was on the repeat sales index comparing energy efficient homes with standard homes. However, as another vital portion of the overall issue, a survey was created to examine how information about a home's energy features is passed during the resale of a property. The survey was designed to examine why owners moved; what features they valued in their old and new homes; what information they found or were given about the new home; and what the sources were for that information. The purpose was to try to determine whether or not energy efficiency was important in the purchase of a home, and which of the key players in a sales transaction (seller, buyer, real estate agent, appraiser, lender, etc.) communicated details about that energy efficiency. Is energy efficiency of a home being communicated effectively and accurately during a sale, and does that correspond with a buyer's interest in energy efficiency? The topic of how information on energy efficiency is transferred during a home sale is a key aspect to the study of the value of energy efficient homes. Knowing how information flows would not isolate the effect on appreciation of energy efficiency in homes. However, it may provide a look into whether communication about efficiency is important in the sale of a home, and may hint at the reasons behind the results of studies like the repeat sales index detailed in this report.

As a federally funded project, the survey had to be designed to limit the burden on survey participants and on record collection and maintenance, so it was designed to be as concise as possible while still collecting the needed information. This was designed to be an electronic survey that would be accessed in a link mailed to the recipient. The time expected for the survey was approximately 5 minutes. The survey also had to be approved by the Office of Management and Budget (OMB). OMB never took action on the survey (neither rejecting nor approving). Therefore, the survey was developed, but never actually sent. Below, the survey is explained question by question.

1. Are you or a member of your household actively employed? (Select one option)

- ○ Yes
- ○ No

This question helps to qualify the results by providing some information about the financial status of the home. A retiree or an independently wealthy buyer may have different motivating factors than a buyer that is working.

2. Do you own your current home? (Select one option)

- ○ Yes
- ○ No

This question was intended to filter out anyone who might receive the survey but might not be the owner of the property. The survey's focus is information transferred during sales, so renters would not be able to give useful answers.

3. How many homes have you purchased, including your current home?

- ○ 1
- ○ 2
- ○ 3
- ○ 4
- ○ 5
- ○ More than 5

This is a demographic helping to qualify the results of the survey. It allows results to be analyzed by whether a buyer is purchasing their first home, a retirement home, or something in between. It also may show how some amount of experience in buying homes changes perspective on what is valued in the home.

4. How far did you move when you purchased your current home?

- ○ 0-25 miles
- ○ 26-100 miles
- ○ 101-500 miles
- ○ Over 500 miles

This question helps to qualify results explaining why a buyer moved from their previous home. A buyer moving over 500 miles is unlikely to be moving that distance because they want better schools or lower utility bills and highly likely to be moving to be closer to work. This helps to explain the reasoning behind a move to a further extent.

5. Why did you want to move from your previous home? (Check all that apply)

- ☐ To downsize
- ☐ To upsize
- ☐ To find a better neighborhood
- ☐ To find better schools
- ☐ To be closer to work
- ☐ To have a lower monthly mortgage
- ☐ To have lower monthly utilities
- ☐ To have lower taxes
- ☐ To have lower home owner association fees
- ☐ To be closer to retail, restaurants, entertainment, and services
- ☐ N/A: have purchased only one home

This question was designed to examine the motivating factors for moving. This can point to what information they may have been looking for in the process of searching for a home. If a buyer is already searching for a home that would lower utility bills, this may tell us something about why they may or may not have gathered information on energy efficient features in a home.

6. How long do you expect to live in your current home?

- ○ Less than 1 year
- ○ 1-3 years
- ○ 3-7 years
- ○ 7-11 years
- ○ 11-20 years
- ○ Greater than 20 years

This question is meant to qualify results by showing how the buyer expects to use their home. Is this a starter home to the buyer, or perhaps an investment home? Or does the buyer consider this the home they will live in for most of their life? The length of time a buyer expects to live in a home sheds some light on what they might value in a home and why.

7. What is your age?

This was a demographic question aimed at qualifying the results of the survey. Being able to sort answers by age may show differences in age demographics in interest in energy efficientcy in a home.

8. In the purchase of your current home, rate the importance of the following features, which ones your previous home contained, and which ones your current home contains. (Check all that apply)					
	Low importance	Medium importance	High importance	Had in previous home	Have in current home
(a) Desirable flooring	☐	☐	☐	☐	☐
(b) Desirable appliances	☐	☐	☐	☐	☐
(c) Desirable exterior finish	☐	☐	☐	☐	☐
(d) Desirable landscaping	☐	☐	☐	☐	☐
(e) Desirable number or size of windows	☐	☐	☐	☐	☐
(f) Energy Star certification	☐	☐	☐	☐	☐
(g) Desirable interior lighting	☐	☐	☐	☐	☐
(h) Low energy cost	☐	☐	☐	☐	☐
(i) Desirable kitchen counters/cabinets	☐	☐	☐	☐	☐
(j) Desirable bathroom fixtures/counters /cabinets	☐	☐	☐	☐	☐

Rating the importance of features in a home allows comparison among the different features and the extent to which the buyer values them. The question also compares the buyer's current and previous home, providing insight into whether a correlation exists between value of a feature and whether the buyer already had it, lost it, or gained it for the first time.

9. Before purchasing your current home, did you collect information on any of the following? (Check all that apply)

- ☐ Utility bills
- ☐ Home warranty
- ☐ Energy Star certification
- ☐ Property taxes
- ☐ Age of heating and cooling equipment
- ☐ Home owner association fees
- ☐ Home owner association rules

This is a central question to the survey. Utility bills show the buyer's interest in the efficiency and cost of operating their home. Energy Star certification examines the buyer's knowledge of the Energy Star brand, the focus of the repeat sales index portion of this study. Home owner association rules and fees, home warranty, and property taxes are all aspects of a home that buyers were likely to examine. Age of heating and cooling equipment could fit into both the energy and standard features of a home. The thrust of this question is to see how often buyers collected information on the energy efficiency of a home compared to non-energy related information.

10. What sources of information did you use when choosing your current home? (Check all that apply)

- ☐ Real estate agents
- ☐ Advertisements by builders and developers
- ☐ Real estate magazines
- ☐ Internet research
- ☐ Word of mouth/friends and acquaintances
- ☐ Classified ads
- ☐ Signs
- ☐ Driving around

This question shows where buyers are looking for information about homes, as well as giving insight into where energy efficiency needs to be advertised if it is going to be communicated to the buyer.

11. During the purchase of your current home was it: (Select one option)

- ◯ Listed by a real estate agent
- ◯ For sale by owner?
- ◯ Other (please specify) ______________

This question was meant to drill down into why a certain source of information may have been used. For example sources of information may change depending on who is advertising the home. If it is advertised by owner, a real estate agent may not be involved in transfer of information at all. This allows results to be qualified by showing how information is transferred in different circumstances.

12. During the purchase process of your current home, who identified any of the following as being desirable features of the home? (Check all that apply)				
	Real estate agent	**Home inspector**	**Previous owner**	**N/A**
(a) Flooring	☐	☐	☐	☐
(b) Appliances	☐	☐	☐	☐
(c) Exterior finishes	☐	☐	☐	☐
(d) Landscaping	☐	☐	☐	☐
(e) Number or size of windows	☐	☐	☐	☐
(f) Energy Star certification	☐	☐	☐	☐
(g) Interior lighting	☐	☐	☐	☐
(h) Low energy cost	☐	☐	☐	☐
(i) Kitchen counters/cabinets	☐	☐	☐	☐
(j) Bathroom fixtures/counters/cabinets	☐	☐	☐	☐

This final question gathers details on what features of the home were advertised favorably and who provided the information. It aims at the core question of whether or not energy efficiency is being advertised or communicated, and how the information is flowing from the major parties involved in the sale of a home.

The survey was designed to add an important qualitative view to help explain the results of the repeat sales index. The index provides statistical quantitative analysis of how appreciation rates compare between energy efficient and standard homes. However, the survey is intended to add another layer to this analysis, offering reasons for the quantitative results, as well as possible avenues to boost the value of energy efficiency in the market. It is designed to examine how information is passed during the sale of a home, who the major players are passing that information, and how that information is valued by the buyer. Any national study comparing appreciation rates of energy efficient homes with those of standard homes would benefit from having an accompanying survey that examines these topics to help qualify the results.

Conclusion

This study fills an important gap in research dealing with energy efficiency in homes and the value of that energy efficiency. It establishes a sound methodology for examining the difference in appreciation rates between energy efficient homes and homes built using standard building methods and features.

The results of this study, limited to only two MSAs, did not find a difference in appreciation rates, but a broader study over a longer period of time is needed to substantiate or disprove these results. This study would use a larger sample, focusing on at least 20 MSAs in different parts of the country with different housing markets and different climates to ensure a robust sample of energy and control homes in each MSA.

Energy Star for Homes would again be a desirable target to act as the energy efficient homes sample, but buy-in from the RESNET and Energy Star communities will be vital in order to do a study of multiple MSAs. The national registry that RESNET is planning to compile of all Energy Star homes will make access to this data easier.

In addition to repeat sales index calculations implemented on a larger scale, a survey similar to the one designed for this study, preferably implemented in the same MSAs as the repeat sales index would help answer important questions. Knowing how information on energy efficiency is transferred, if at all, during the process of a home sale may help explain or qualify the results of the more quantitative repeat sales index.

This robust sample, in combination with a survey, should result in a more definite answer, which would give important insight into the value of energy efficiency in homes, as well as provide guidance for efforts to educate home buyers about energy efficiency.

Appendix A

Date

Name
Organization
Address
City, State, Zip

Dear HERS Provider:

The Partnership for Advancing Technology in Housing (PATH - a division of U.S. HUD) is sponsoring a study to investigate how energy efficient (EE) homes appreciate over time as compared to baseline homes. Because this study intends to be the most comprehensive study on the valuation of EE homes to date, PATH has enlisted the help of EPA's Energy Star Homes, and RESNET, among other organizations, to supply project oversight. Collectively, we are now looking to HERS providers for your assistance in sharing the data necessary to make this study a success.

The first step in this study is creation of a database of EE homes built within a few targeted metropolitan statistical areas. As the most recognized national metric for EE homes has been a HERS score of 86/Energy Star certification, we are contacting HERS providers to assist in this study by sharing addresses of homes that have been rated to satisfy HERS 86/Energy Star criteria.

The only information that we are looking for is the address of a home, the HERS score, and the date of the rating. We are only interested in ratings conducted on homes that were new at the time of the rating. HERS raters will be kept anonymous, and HERS providers will be thanked through a reference in the final report (unless they request to remain anonymous). Addresses will not be shared with the public and will not be used for any other purpose than this study unless express permission is given by the provider. The overall results of the study will be made publicly available upon completion of the project.

A representative from Newport Partners, a building consulting firm contracted to manage the project, will be contacting you soon to enlist your support. Thanks in advance for your participation in this important study and for your contribution towards building energy efficient homes.

Sincerely,

Mike Moore, P.E.
Project Manager
Newport Partners

Sam Rashkin
National Director
U.S. EPA Energy Star Homes

Dr. Regina Gray
Social Science Analyst
Partnership for Advancing Technology in Housing

Steve Baden
Executive Director
RESNET

3760 Tanglewood Ln, Davidsonville, MD 21035
301.889.0017 • Fax: 301.889.0019 • www.newportpartnersllc.com

Appendix B

Original Study Methodology

Under the original design of this study, Newport planned to conduct in-depth case studies of houses around the country that included energy efficient features, in addition to case studies that would be performed on control houses similar to the energy efficient houses but without the energy features. These case studies would include information about the home such as energy features, size, location; sales history information such as original price and price at resale; and details about how information was passed between the main parties involved including seller, buyer, real estate agent, appraiser, lender, etc. This last item was intended to be the main focus of the study and would include in-depth interviews of all parties involved that could be contacted and were willing to participate. The idea was to see how information about the energy efficiency in a home was passed from party to party, who knew about the features, and who placed a priority on them. Newport then planned to compare the case studies of energy efficient homes to control homes and examine the differences in appreciation between the two. This methodology was abandoned because of major obstacles in favor of an alternative combination of a targeted quantitative analysis in two Metropolitan statistical areas and a qualitative survey.

Below is a description of the work that was done under the original methodology, the work planned, and the reasons for changing the approach of the study.

Work Completed under the Original Methodology

Newport first established a set of sampling criteria with extensive contributions from its subcontractor, Abt Associates (Abt). The sampling criteria step was the only major portion of the study that was completed before the decision was made to adjust methodologies and take the study in a new direction. The sampling criteria created included a plan to perform a total of 8-12 case studies of energy efficient homes, as well as the control group. The criteria gave priority to location, size and structure type, year built, and the extent of energy features. Below the reasoning is laid out for the sampling priorities chosen.

> ***Location*** was considered important in the sampling criteria in that the representativeness for the set of case studies would be enhanced by selecting locations for the case studies according to the frequency of energy efficient homes around the country. To the extent that the results from the case studies agree with one another, that would bolster confidence in the general conclusions. However, possible differences in outcomes could be due to many factors – many more than Newport would be able to control for with such a small sample. Therefore, drawing a sample of case studies distributed across the country would be indicative, not definitive, about the capitalization of energy efficiency.
>
> ***Size and structure type*** was set in the sampling criteria in order to have closely matched energy usage except for the energy efficient features and so that future buyers would consider the properties as viable alternatives. A 2000 square foot home built as a detached house is likely to appeal to a different buyer than an 800 square foot attached house. The differences in energy efficiency would be dwarfed by the larger differences in family size and income of the buyers. Newport decided to focus their sampling criteria on homes between 2,000 and 3,000 sq. ft. Newport also wanted to minimize any unusual stylistic characteristics which could make it harder to match to other houses at original sale or impact the market appeal 10 to 20 years later. For these reasons, Newport planned to exclude unconventional passive solar designs which may have excellent energy efficiency, but limit the resale market to energy conservation enthusiasts. Because of the limited sample size, Newport planned to limit the structure variation to single-family detached houses with 3 to 4 bedrooms and 1.5 to 2 bathrooms.
>
> ***Year built and resold*** in the sampling criteria was a primary concern. Newport set the criteria so that the energy efficient houses and their non-energy efficient companions would be built within the same 12 to 18

months. Most developments meet this criterion. And local housing markets usually evolve gradually enough so that the original purchase price reflects the energy efficiency premium rather than a shift between buyer and seller markets. Newport set a secondary criterion that the houses would be built in the 1990-1996 period to allow for about 10 years of seasoning before resale. The seasoning would be a fundamental test as to whether the energy savings is permanently capitalized into the value of the house. It is possible that new high performance homes carry a market premium which fades fairly quickly over 5 to 7 years. After that initial period, but well before system replacement, Newport wanted to know if the energy efficient features continue to boost resale value. This mid-range in years would be important to original buyers because it is closer to the typical holding period for homeowners and before many energy efficient systems have paid for themselves in energy savings.

Ideally, the energy efficient home and companion home would be sold within 12 to 18 months of one another. Newport suspected that this would be a challenging criterion to meet. One approach would be to start with many companion properties and then select the few with the closest resale dates (before and after the energy efficient house resale). The comparison price could be imputed from averaging the resale values of the non-energy efficient houses.

Extent of energy efficient features in the sampling criteria needed to be important enough to generate significant savings in utility bills such that the seller and their real estate agent promote this feature in the sales pitch. A low-cost feature may be advantageous to the original buyer in quickly repaying for the investment, but have little impact on the resale value because the seller's agents and buyer take no notice of it. Newport settled on a sampling criterion of an energy efficient premium of $20,000 to $50,000 to separate energy efficient houses significantly from the control group. Newport set overall price point targeted for the sample as the median home value for that area before the energy efficient premium. Newport set the control group price point would be the median house value. If the case studies failed to detect that size of effect, then it is likely smaller effects would also not be noticed. In either case, it would take a larger study to precisely measure the appreciation rate, but some case studies finding some evidence of impact may be a stronger basis for further study.

There are many more dimensions along which house values and energy consumption can vary. For example, ***number of occupants*** will certainly affect the utility bills and energy savings. Ideally the energy efficient and non-energy efficient units in the sample would have the same number of occupants, but given the many higher priorities the most Newport could plan to do is exclude outliers. ***Remodeling and additions*** can greatly enhance the size and value of a house. The ideal comparison houses would have average maintenance and no additions, but Newport would likely only be able to find out about the more extreme cases based on the square foot size or number of rooms advertised at resale. A subset of remodeling could be adding more energy efficient features. Original buyers of high performance houses may be avid conservationists and keep adding more insulation or solar panels. It might be that these changes may enhance the energy savings and boost the resale values beyond the efficiency premium at the original sale. For the same reason as excluding retrofits, Newport wanted to avoid major energy efficiency enhancements that could bias the measure of capitalization.

Finally, Newport developed three possible approaches to identifying sample energy efficient homes and control homes.

- **Approach 1**: One development in which the same builder built both the energy efficient and baseline homes. This was the best option because it ensured that location, model, and building practices would all be similar. The difference would be the energy efficient features, providing the closest sample and the best way to focus on the effects of energy efficiency.

- **Approach 2**: Two nearby developments in which the same builder built both the energy efficient and baseline homes. This was the next best choice because using the same builder would still likely result in similar models and building practices. A nearby development would ensure that houses being compared would be in the same market.

- **Approach 3**: High performance home with comparable baseline homes nearby, but not necessarily built by the same builder. This was the least favorable option because it introduces the variable of differences between builders, their techniques, and their marketing.

As mentioned above, developing the sampling criteria was the only major step completed before abandoning the original methodology because the sampling criteria could not be met.

Work Started Under the Original Methodology but Not Completed

After establishing sampling criteria, Newport began to search for suitable case study locations by approaching builders and green/energy efficient building programs to help identify energy efficient homes and similar control homes. This was the step that caused Newport to abandon the planned methodology and explore alternatives. Difficulties in collecting enough information on specific sites, as well as cost prohibitive amounts of labor necessary to overcome these obstacles pointed the project in a new direction. These barriers are described further at the end of the "Original Methodology" section.

The next step planned was to develop an interview protocol and scripts that would be tested on a pilot case study site. It would be used to guide the interview process with owners and professional parties involved in the sale of the house. The Office of Management and Budget (OMB) would need to approve these instruments before use in the study as part of the Paperwork Reduction Act. This step was started but never completed as methodology was changed during development of protocols and scripts.

Further Work Planned Under Original Methodology but Not Started

Newport planned to begin outreach, after testing and OMB approval, by sending a letter of introduction explaining the study to the current home owner and then try to recruit the other involved parties in that home's sale after gaining owner participation. Newport also planned to include a call back number and email address for those wishing to respond on their own initiative, expecting approximately 20% of respondents will self-identify this way. An incentive of a Home Depot gift certificate would be included for the respondents.

Newport's next step was to be developing sales history assessment extraction forms that would standardize the information being collected from publically available sources on each selected property. Newport would then conduct interviews with stakeholders and homeowners of selected properties using the interview protocol approved by OMB. The interviewees would be selected from the interview participant list of respondents collected through outreach. To ensure good quality interviews Newport planned to do the following:

- Prior to conducting the interviews, review the relevant information from the sales history assessments.
- Be familiar with the interview protocol and practice conducting interviews at an internal staff training session.
- After each interview, Newport would hold a de-briefing to discuss emerging themes and fine-tune the interview protocol.

Newport would collect the sale history and interview results from each case study and compare the case studies within each group (energy efficient homes and control homes). Newport planned to summarize themes and patterns within each group. Newport would then compare across groups to find what patterns are similar and different between the energy efficient and control homes. In the final report, any proprietary information would be removed and participants would remain anonymous. The study was designed to provide national qualitative analysis of energy efficient homes; how information about that home is transferred and by whom; and how this compares to standard homes. Although the results would not be statistically significant, they were intended to provide a rich, individual story for each case study that could be used to direct further, more quantitative research on the topic.

Difficulties and Barriers found in the Original Methodology

Newport ran into major barriers trying to identify homes and communities to use as case studies in this project, including:

- Lack of builder records
- Lack of builder willingness to participate
- Lack of national database of Energy Star homes
- Lack of time or funds to make numerous site visits to determine whether homes identified through other means would be suitable case studies.

The first problem was that many of the builders that we contacted did not keep detailed records of the communities they had done or the addresses of the homes. Either they did not have records at all, or they did not have detailed records of the features of these homes. If a builder had been able to provide us with addresses or communities to visit, it would have taken a trip to the house or community just to determine whether it would be suitable for a case study. Site visits were expected to collect data for the case studies, but not to simply determine whether a site would work or not. The next barrier that Newport encountered was that builders were worried about privacy concerns for their clients. Because the case study would involve attempting to contact the current owner and previous owner, some builders balked at the idea of having their clients inconvenienced for a study. Builders also were hesitant to provide information to a study that would potentially be comparing their efficient products to their standard products, or to products of other builders. They did not want former clients contacting them and asking them why they were sold an "inferior" product. They did not want analysis done that would show one of their products performed better or worse during resale than another of their products. They did not want to have their products used as the control group for a similar builder using energy efficient features because it would make their product appear inferior to that of their competitors.

Because a significant amount of time was spent trying to find builders willing to identify possible case study subjects, but had not achieved any progress, Newport suggested developing an alternative methodology that would use a repeat sales index to compare appreciation rates of energy efficient homes to those of standard homes in the same Metropolitan Statistical Area (MSA) and built at the same time. This would be a quantitative approach, rather than the more qualitative, interview intensive approach of conducting case studies. There were fewer barriers to privacy collecting publically available data on a larger scale, greater ability to identify addresses when the focus wasn't houses by the same builder, or in the same or neighboring communities, and greater statistical significance than case studies would have provided. This would still achieve a comparison of the value of energy efficient homes compared to standard homes. In order to keep some of the feel of the original proposal, Newport suggested a survey that would obtain information about what was important to buyers in the purchase of a home, how they found out information about the home, and who was involved in passing this information.

Sources Cited

Ahluwalia (2008) “Consumer Preference Survey 2007-2008”, National Association of Home Builders, Economics Group

Bailey, M., Muth, R., and Nourse, H. (1963), "A regression method for real estate price index construction," *Journal of the American Statistical Association*, 58, 933-942.

Black (2003) “Financial Payback On California Residential Solar Electric Systems”

Black (2004) “Financial Payback On California Residential Solar Electric Systems After The State Rebates Are Gone”

Calhoun “OFHEO House Price Indexes: HPI Technical Description.”
Available at: www.ofheo.gov/Media/Archive/house/hpi_tech.pdf.

Case and Shiller (September/October, 1987), "Prices of single-family homes since 1970: new indexes for four cities," *New England Economic Review*, 45-56.

Cortright, (May, 2008) "Driven to the Brink: How the Gas Price Spike Popped the Housing Bubble and Devalued the Suburbs" *CEOs for Cities*

Census Bureau method to construct price indices for new homes sold or under construction.
See http://www.census.gov/const/www/constpriceindex.html

Dacquisto, Emrath, Laquatra, and Laitner (December, 2001) “The Value of Energy Efficiency in Housing: Review and Analysis of The Literature

Energy Star for Homes website.
See www.energystar.gov/homes

Energy Star Milestones.
See http://www.energystar.gov/index.cfm?c=about.ab_milestones

Goetzmann, and Spiegel, (1995), "Non-temporal components of real estate appreciation," *Review of Economics and Statistics*, 77 (1), 199-206

Nevin and Watson (October, 1998) “Evidence of Rational Market Values for Home Energy Efficiency”, *The Appraisal Journal*

Office of Management and Budget (2007) “Bulletin No. 08-01”

OFHEO News Release (March 1, 2007), pp 8-12 "The Market Slowdown and Home Prices in the Suburbs and 'Exurbs'"

RESNET 2006 Mortgage Industry National Home Energy Rating Systems Standard

Stiff (May 27, 2008), "Housing Bubbles Collapse Inward" *Fiserve Lending Solutions*