Sawtooth Software

RESEARCH PAPER SERIES

Testing Adaptive CBC: Shorter Questionnaires and BYO vs. "Most Likelies"

Bryan K. Orme & Richard M. Johnson,
Sawtooth Software, Inc.

530 W. Fir St.
Sequim, WA 98382
(360) 681-2300
www.sawtoothsoftware.com

Testing Adaptive CBC:
Shorter Questionnaires and BYO vs. "Most Likelies"

Bryan Orme and Rich Johnson, Sawtooth Software, Inc.

Background:

This research reports a third successful application of a new form of Adaptive Choice-Based Conjoint (ACBC) as first proposed by Johnson and Orme (2007). We'll assume the reader is familiar with our 2007 paper (available at www.sawtoothsoftware.com/techpap.shtml), so we can focus on the new findings and developments presented here.

The first application of this ACBC method involved laptop computers (Johnson and Orme, 2007), the second application was an actual industry study for a type of motorized recreational equipment (also reported in Johnson and Orme, 2007), and this current study involved home purchases.

In the first two applications, the ACBC survey included three core sections:

1. **BYO (Configurator) Section**, in which respondents specify their preferred product, subject to prices attached to each level. As they configure their preferred option, the total price is updated and shown.

2. **Screening Section**, in which respondents specify whether they would or would not consider various products (in full-profile). The products were constructed using a near-orthogonal plan, to be near-neighbors to the BYO-specified product.

3. **Choice Tasks Section**, in which respondents choose among products which they indicated in the Screening Task that they would consider. These products are shown in a series of standard CBC questions arranged in a round-robin tournament, until an overall winning product is identified.

Key results from the first two tests were as follows:

- Respondents found the ACBC interview more realistic (presenting more realistic and relevant products) and more engaging than standard CBC.
- Respondents took double or triple the time to complete the ACBC questions as respondents who received standard CBC questionnaires. Despite the increased length, they reported equal satisfaction with the survey.
- The aggregate utilities for ACBC and standard CBC were correlated at 0.9 or better.
- ACBC provided more accurate predictions (than standard CBC) of CBC-looking holdout tasks that were customized (at the individual level) to include winning concepts from previous CBC holdout tasks. This was not necessarily expected, as methods bias would have favored CBC-developed utilities in predicting CBC-looking holdouts.

- ACBC provided more accurate predictions of choice shares for holdout respondents on CBC-looking holdout choice tasks, when including the respondents who took the most time (and were presumably being the most careful) in answering their holdout tasks.

When we presented these findings at the 2007 Sawtooth Software Conference, a couple of concerns were raised by members of the audience:

- Some product categories don't naturally work well within a BYO (Configurator) task. Furthermore, prices attached to attribute levels might be difficult to determine and agree upon with the client.
- The interview was significantly longer (double or triple the length of the CBC interview). In the Laptop study, ACBC averaged 11.6 minutes to CBC's 5.4 minutes. In the Recreational Equipment study, ACBC averaged about 15 minutes compared to about 5 minutes for CBC.

Interestingly enough, also at the 2007 Sawtooth Software Conference, a similar approach to our ACBC was introduced by Gaskin *et al.*, as a cooperative effort between AMS, INSEAD and MIT (Gaskin *et al.* 2007). That paper was developed independently of our work, and also included the idea of a consideration phase (to identify non-compensatory rules) followed by a customized choice-tasks phase.

Goals of the Current Research:

In this most recent test, we investigated the following issues:

- Can similar results (predictions meeting or exceeding CBC for CBC-looking holdout tasks) be obtained after shortening the ACBC interview?
- Can we substitute a "**Most Likelies**" section for the **BYO** section? A most-likelies section would ask respondents simply to indicate which level for each attribute (except price) they would most likely have in the product they would buy. No prices were shown attached to these levels.
- Did including a graphic of an attractive interviewer improve respondents' perceptions of the interview or quality of responding?

We also made a small methodological improvement in the way adaptive designs were built on-the-fly.

Description of Attributes and Questionnaire:

Based on our first two tests of ACBC, we felt that this new methodology would have advantages over traditional CBC research for studies involving about five or more attributes for high-involvement purchases. We decided that home purchasing definitely fit this description, and used the following attribute list:

Table 1
Attribute List

1. Size of Home:
1400 Square Feet
1750 Square Feet
2250 Square Feet
3000 Square Feet
4000 Square Feet

2. Eco-Friendly Features:
Standard construction & efficiency
Built green certification & superior efficiency

3. Bedroom/Bathroom Configuration:
2 bedrooms, 1.5 bath
3 bedrooms, 2.5 baths
3 bedrooms, 3 baths
4 bedrooms, 3.5 baths
5 bedrooms, 3.5 baths

4. Master Bath:
Small master bath (integrated shower/tub), small walk-in closet
Medium master bath (separate shower and tub), one walk-in closet
Large master bath (separate shower & deep soaking tub), two walk-in closets

5. Garage:
2-car garage
3-car garage

6. Lot:
Small lot (0.2 acres)
Medium lot (0.35 acres)
Large lot (0.5 acres)

7. Schools:
Local schools ranked in bottom third of nation
Local schools ranked in middle third of nation
Local schools ranked in top third of nation

8. Finish:
Standard finish package (floors, countertops, millwork, fixtures)
Upgraded finish package (floors, countertops, millwork, fixtures)
Premium finish package (floors, countertops, millwork, fixtures)

9. Age of Home:
Home built in 1980
Home built in 1990
Home built in 2000
New home

10. Price (conditional, based on US region and square footage of home):
-25%
-10%
+10%
+25%

Conditional (Summed) Pricing:

One of the challenges in conducting research on homes is that prices can vary dramatically based on the region in the US and the city where the respondent would be purchasing. We conducted a preliminary search on the internet for published statistics on the average home prices for different square footage homes based on different regions of the country. At the beginning of the interview, we asked respondents the following question:

During this survey, we'd like you to imagine that for some reason you had to leave your current home and purchase a different home. We'll be asking to you evaluate homes with different features, to learn about what types of homes you prefer.

Different cities can have quite different prices for homes. So that we can show you prices appropriate for your area of the country, please select the price tier that best applies to the area where you would buy a home.

If you were to purchase a home, would this home be located in:

- O **Premium cost area (such as San Francisco bay area; Long Island, NY; Honolulu, HI) Median home price about $700K**
- O **High cost area (such as Seattle, WA; Boulder, CO; Orlando, FL) Median home price about $300K**
- O **Medium-low cost area (such as Columbus, OH; Dallas, TX; Salt Lake City, UT; Durham, NC) Median home price about $200K**
- O **Low cost area (such as Buffalo, NY; Wichita, KS; Pittsburgh, PA; Oklahoma City, OK) Median home price about $150K**

Based on the respondent's selection, we multiplied all prices shown to the respondent by the following multipliers (and then rounded to the nearest $250):

Premium Cost Area:	2.25
High Cost Area:	1.25
Medium-Low Cost Area:	0.80
Low Cost Area:	0.50

For both the standard CBC and ACBC interviews, we also made the prices shown conditional on the square footage of the home. The base prices (prior to multiplying by the regional cost index above) were as follows:

1400 Square Feet ($210,000)
1750 Square Feet ($255,000)
2250 Square Feet ($315,000)
3000 Square Feet ($405,000)
4000 Square Feet ($520,000)

Thus, for a respondent selecting a High Cost Area, a 2250 square foot home at the low price point (-25%) would be presented in the choice tasks at a price of 315,000 x 1.25 x 0.75 = $295,250.

For the ACBC questionnaire, we employed conditional (summed) prices. In addition to prices being conditional on square footage (as with the CBC questionnaire), we computed prices for the homes based on additional features:

Table 2
Level Prices

Eco-Friendly Features:
Standard construction & efficiency
Built green certification & superior efficiency (Add $15,000)

Bedroom/Bathroom Configuration:
2 bedrooms, 1.5 bath (subtract $15,000)
3 bedrooms, 2.5 baths (subtract $5,000)
3 bedrooms, 3 baths
4 bedrooms, 3.5 baths (Add $10,000)
5 bedrooms, 3.5 baths (Add $10,000)

Master Bath:
Small master bath (integrated shower/tub), small walk-in closet (subtract $10,000)
Medium master bath (separate shower and tub), one walk-in closet
Large master bath (separate shower & deep soaking tub), two walk-in closets (add $10,000)

Garage:
2-car garage
3-car garage (add $10,000)

Lot:
Small lot (0.2 acres) (subtract $20,000)
Medium lot (0.35 acres) (add $0)
Large lot (0.5 acres) (add $20,000)

Schools:
Local schools ranked in bottom third of nation (subtract $30,000)
Local schools ranked in middle third of nation
Local schools ranked in top third of nation (add $30,000)

Finish:
Standard finish package (floors, countertops, millwork, fixtures)
Upgraded finish package (floors, countertops, millwork, fixtures) (Add $15,000)
Premium finish package (floors, countertops, millwork, fixtures) (Add $30,000)

Age of Home:
Home built in 1980 (subtract $15,000)
Home built in 1990
Home built in 2000 (add $15,000)
New home (add $30,000)

And, rather than vary the final prices by -25% to +25% as with the standard CBC design, we chose to vary the summed prices by -15% to +15%[1]. This had the positive effect of making prices more correlated with features for ACBC (than for CBC), but this also brought negative consequences of greater multicolinearity in the independent variable matrix.

We also included a few prohibitions (for both CBC and ACBC), to ensure that the most unrealistic home combinations would never be shown. The following combinations were prohibited:

1400 sq. ft. *with* 4 bedrooms 3.5 baths
1400 sq. ft. *with* 5 bedrooms 3.5 baths
1750 sq. ft. *with* 5 bedrooms 3.5 baths
3000 sq. ft. *with* 2 bedrooms 1.5 baths
4000 sq. ft. *with* 2 bedrooms 1.5 baths
Built green certification & superior efficiency *with* Home built in 1980

Prior to fielding, we tested the design efficiency of both CBC and ACBC interviews and determined that the standard errors of all estimated effects (including the coefficient for price) were reasonable.

Experimental Plan and Data Collection:

The attribute list for this current research involves 10 attributes, 31 total levels (counting price as a single linear parameter), and 22 parameters to be estimated. A key decision to make when designing an ACBC questionnaire is how many product concepts to show in the Screener Section. This has the greatest impact on the total length of the ACBC questionnaire.

The designs for the current as well as the previous two ACBC projects are as follows:

Table 3

	Total Attributes	Total Levels	Parameters to Estimate	Concepts Presented
Laptops	10	33	24	40
Recreational Equipment	8	30	22	36
Homes	10	31	22	Either 24 or 32

In the current test for home purchases, we decided to assign respondents randomly to receive either 24 or 32 concepts to evaluate in the Screening Section. This represents a significant decrease in the size of the design relative to the previous two tests.

[1] We varied summed price for ACBC by a smaller random price variation than CBC so that the absolute range of prices was more similar. Both methods used the same base prices based on square footage, but ACBC additionally incorporated level-based prices for the other features. If we used the same price variation across the methods, homes in the ACBC interview would have varied much more widely in terms of absolute total price.

Data were collected in Feburary, 2008 using Opinion Outpost, from Western Wats. The sample included US homeowners, age 25+. Respondents were randomly assigned to one of four questionnaires:

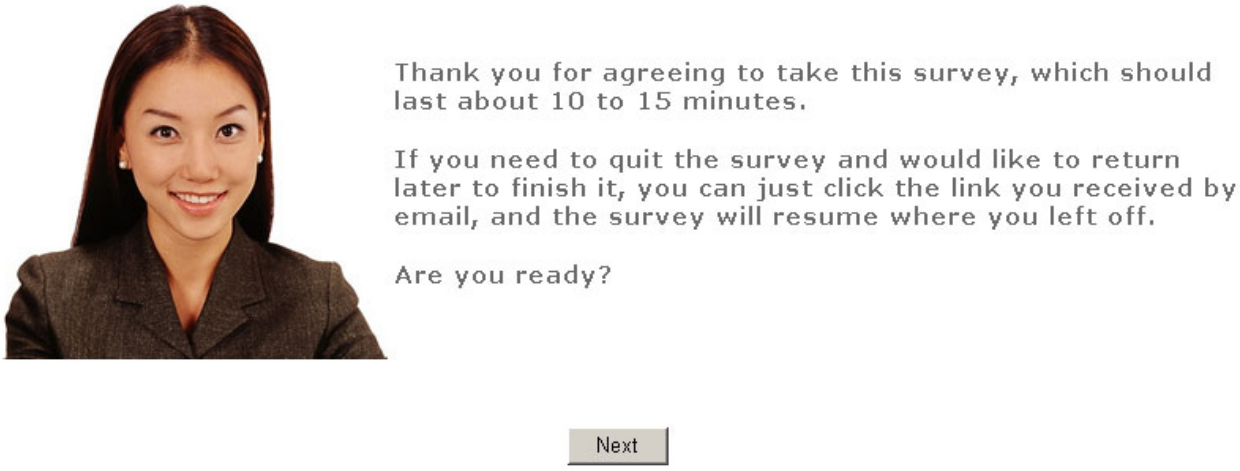
1. ACBC survey, 32 concepts (BYO + Screening + Choice Tasks)
2. ACBC survey, 24 concepts (BYO + Screening + Choice Tasks)
3. ACBC survey, 32 concepts (Most Likelies + Screening + Choice Tasks)
4. CBC survey, 18 tasks, 4 concepts per task + "None[2]" alternative. (See Appendix A for example screen layout.)

A functioning version of questionnaire version 2 (without the holdout questions and qualitative evaluation questions) is available at www.sawtoothsoftware.com/test/acbc/hou1logn.htm.

Respondents were deleted if they answered either too quickly or too slowly (indicating that they had probably interrupted their surveys and resumed later). After deleting the slowest 7.5% and fastest 7.5%, the following sample sizes remained by cell:

Cell 1, n=299
Cell 2, n=303
Cell 3, n=295
Cell 4, n=314

In the previous laptop study, we had included a graphic of an attractive interviewer (see below) who periodically appeared on the screen, and the script was written in first-person singular, as if she were having a conversation with the respondent:



[2] The percentage of choices of None for Cell 4 was 17%, in line with many other CBC projects we've examined.

The question naturally arises whether this personalization made a difference in how respondents reacted to the survey. So, for this home buying questionnaire, respondents were randomly assigned one of two conditions: receive graphic or not.

The BYO Section:

Respondents in design cells 1 and 2 for ACBC received a BYO section. Prices were customized based on the region of the country (price index) chosen by the respondent. See Table 2 for the prices used[3].

Please assume you were in the market to buy a home. If these were the available options, what home would you be most likely to buy?

As you select different options, the total price of the home will change. When you have finished choosing the home that suits you best, click the Next button to continue.

Feature	Select Option	Option Price
Size of Home	1750 Square Feet $318,750	$ 318,750
Eco-Friendly Features	Standard construction & efficiency	$ 0
Bedroom/Bathroom Configuration	3 bedrooms, 2.5 baths (subtract $6,250)	$ -6,250
Master Bath	Medium master bath (separate shower and tub), one walk-in closet	$ 0
Garage	2-car garage	$ 0
Lot	Large lot (0.5 acres) (+ $25,000)	$ 25,000
Schools	Select Option	$
Finish	Select Option	$
Age of Home	Select Option	$
	Total:	$ 337,500

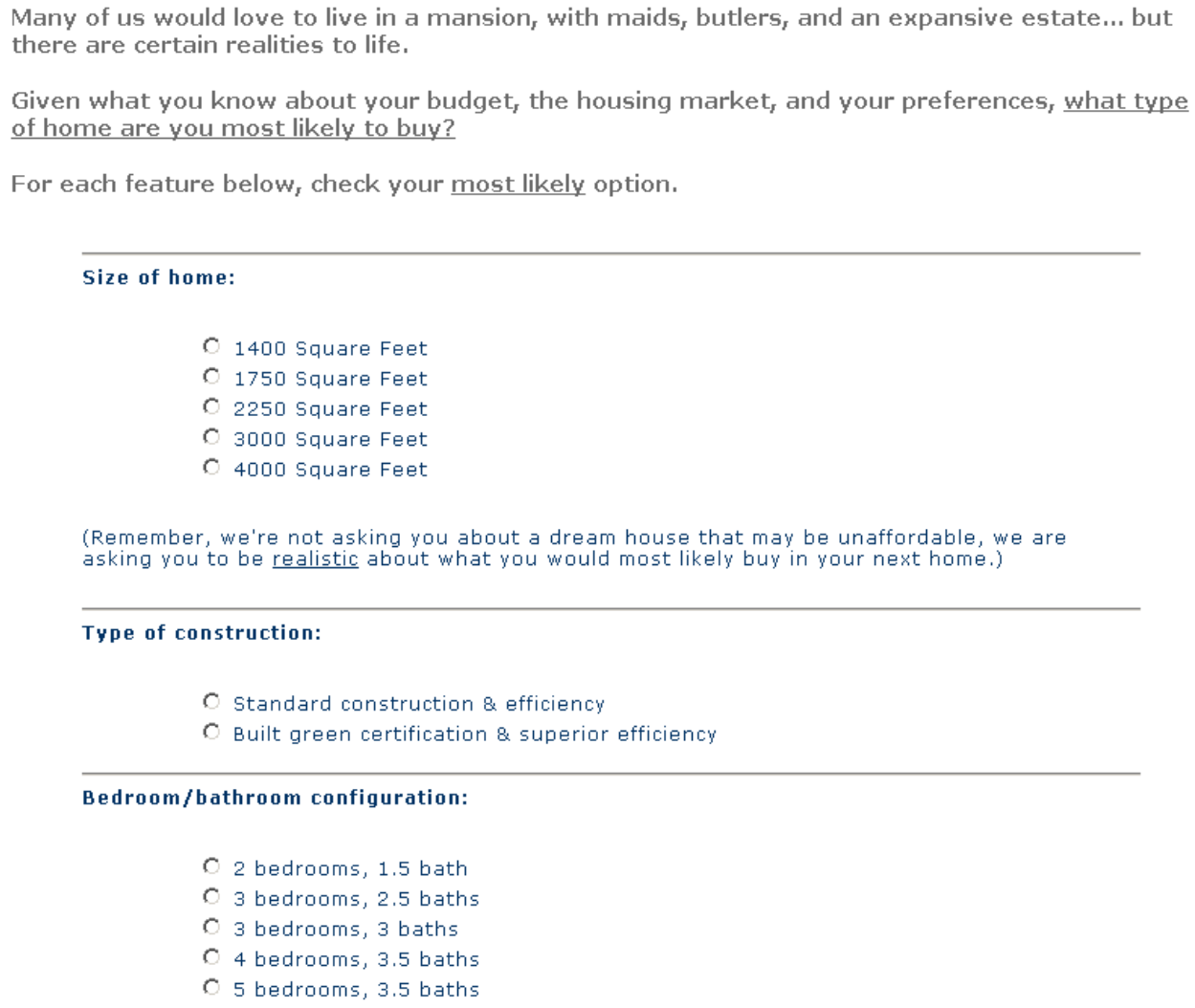
The "Most Likelies" Section:

ACBC respondents in design cell 3 received a "Most Likelies" section rather than the BYO section. Some people have commented that a BYO task doesn't make sense for some product

[3] Some researchers may question whether the specific prices attached to the levels matter much to the resulting part-worth utilities for ACBC. If so, this would lead to significant challenge in developing ACBC questionnaires. Our three tests have shown that level-based prices we selected resulted in overall part-worth utilities very closely resembling CBC utilities (correlation 0.9 or better, based on aggregate results) and of course successfully predicting CBC-looking holdouts. Furthermore, in the Recreational Equipment study, we manipulated (significantly) the prices attached to levels for a few of the attributes. After partialing out the effect of price (the standard ACBC price estimation with a linear term), the part-worth utilities for the attributes where price levels varied were quite similar, irrespective of the price manipulation. This experiment and finding was not reported in Johnson and Orme 2007.

categories. Yet, ACBC relies on learning early on what levels are most relevant to the respondent's choice so that it can design relevant products to evaluate within the Screening Section. The idea of asking about "most likely" levels is not new: it was included as an optional section within early versions of Sawtooth Software's DOS-based ACA program.

Here is an example of our Most Likelies section, for the first few attributes (though all attributes, except price, were included in this section):



The Most Likelies section asks respondents to indicate what they would actually (realistically) buy, keeping in mind their budget and requirements. It also assumes that respondents understand that certain levels would carry premium prices, such as higher square-footage homes costing significantly more. Our analysis of the counts (the percent of times respondents selected each level in either the Most Likelies section or the BYO section) suggests that respondents generally made very similar selections, irrespective of whether the prices per level were shown (as in BYO) or not. See Appendix B for the comparative counts.

We should note that the analysis of Most Likelies is different from the BYO section, and special considerations should be made. We'll discuss this in further detail later in this paper.

The Screening Section:

Respondents receiving ACBC surveys completed a Screening Section, wherein they evaluated homes that were constructed as near-neighbors to their previously specified BYO or Most Likely product. Respondents saw four homes on each screen. Respondents in Cells 1 and 3 received 8 screens of 4 homes each (32 total homes). Respondents in Cell 2 received 6 screens of 4 products each (24 total homes)

Here are a few homes you might like. Do any of these look like they are possibilities? It's helpful if you can keep about half of them for further consideration. But, it's up to you.

(1 of 8)

1750 Square Feet Built green certification & superior efficiency 3 bedrooms, 2.5 baths Medium master bath (separate shower and tub), one walk-in closet 2-car garage Medium lot (0.35 acres) Local schools ranked in middle third of nation Upgraded finish package (floors, countertops, millwork, fixtures) Home built in 2000 $350,250	2250 Square Feet Standard construction & efficiency 3 bedrooms, 3 baths Medium master bath (separate shower and tub), one walk-in closet 3-car garage Small lot (0.2 acres) Local schools ranked in top third of nation Standard finish package (floors, countertops, millwork, fixtures) Home built in 1990 $481,500	2250 Square Feet Standard construction & efficiency 4 bedrooms, 3.5 baths Small master bath (integrated shower/tub), small walk-in closet 2-car garage Large lot (0.5 acres) Local schools ranked in middle third of nation Upgraded finish package (floors, countertops, millwork, fixtures) Home built in 1990 $459,500	2250 Square Feet Standard construction & efficiency 3 bedrooms, 3 baths Medium master bath (separate shower and tub), one walk-in closet 2-car garage Medium lot (0.35 acres) Local schools ranked in middle third of nation Premium finish package (floors, countertops, millwork, fixtures) Home built in 2000 $517,500
○ A possibility ○ Won't work for me	○ A possibility ○ Won't work for me	○ A possibility ○ Won't work for me	○ A possibility ○ Won't work for me

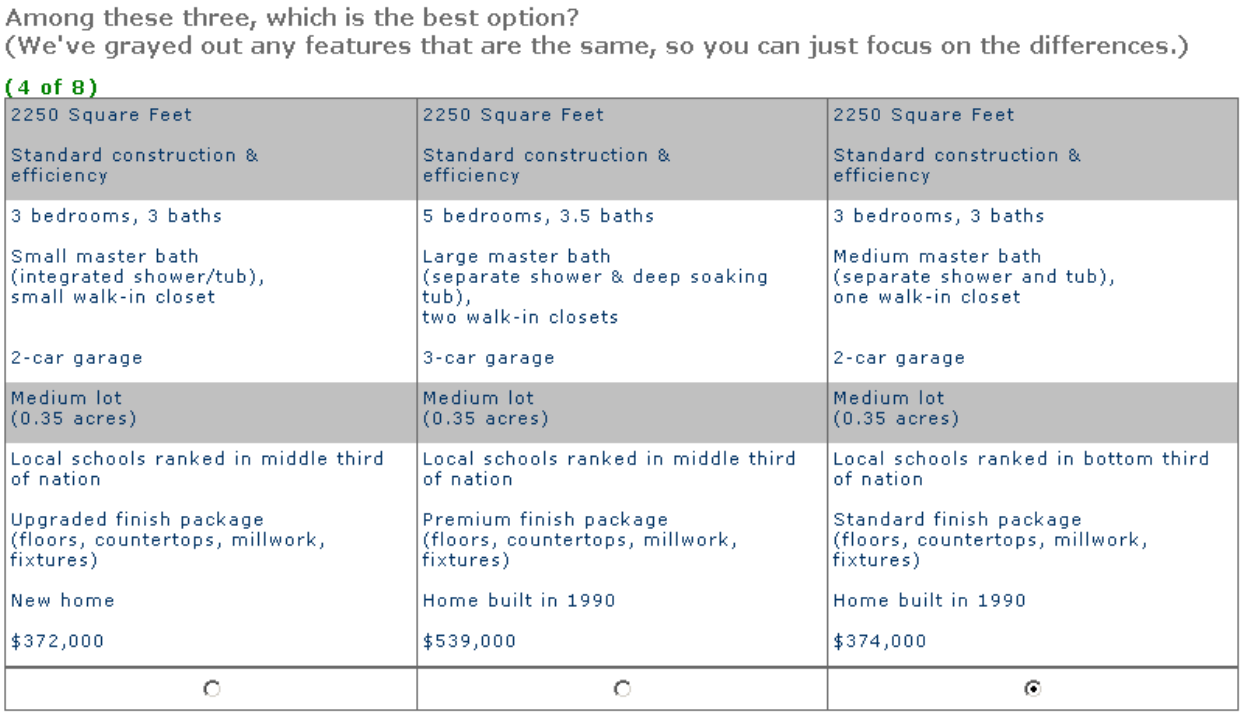
After respondents evaluated three screens of homes, they were asked regarding any "Must Have" or "Unacceptable" levels after each subsequent screen. We scanned the previous choices and inferred what cutoff rules the respondent might be employing, and asked the respondent to confirm the cutoff rule or deny using such a rule. If respondents confirmed a "must-have" rule, such as "I need a home that is at least 2250 square feet," then all homes shown to the respondent in later screens would have at least 2250 square feet. This process was described in more detail in Johnson and Orme (2007).

As mentioned earlier, we modified our approach for generating the pool of concepts to screen. In previous tests, we generated a fixed number of tasks (such as 40, in the case of the Laptop study). If respondents confirmed cut-off rules, any concepts not yet seen that contained rejected levels were automatically marked as "won't work for me." Thus, a respondent might only need to see and evaluate, for example, 35 of the 40 cards in this section.

With this current study, we decided to fix the number of concepts that the respondent would explicitly judge. Thus, respondents in the 32-card cell of our design always evaluated 8 screens with 4 cards each. When cut-off rules were established that would eliminate cards not-yet-seen from consideration, replacement cards were generated (meeting the cut-off criteria) to replace the now-eliminated cards. Thus, respondents in the 32-card cell might have reflected in their data that they "evaluated," for example, 40 cards, even though they only explicitly evaluated 32.

The Choice Tasks Section:

This section uses homes selected as "a possibility" in the Screening Section. The homes are shown three at a time, formatted as CBC tasks, in a "choice tournament." Winning homes are carried forward to subsequent choice tasks, and the process repeats until an overall winning home is selected. Any attributes that are the same across the three homes within a task are "grayed out."



Among these three, which is the best option?
(We've grayed out any features that are the same, so you can just focus on the differences.)

(4 of 8)

2250 Square Feet Standard construction & efficiency	2250 Square Feet Standard construction & efficiency	2250 Square Feet Standard construction & efficiency
3 bedrooms, 3 baths Small master bath (integrated shower/tub), small walk-in closet 2-car garage	5 bedrooms, 3.5 baths Large master bath (separate shower & deep soaking tub), two walk-in closets 3-car garage	3 bedrooms, 3 baths Medium master bath (separate shower and tub), one walk-in closet 2-car garage
Medium lot (0.35 acres)	Medium lot (0.35 acres)	Medium lot (0.35 acres)
Local schools ranked in middle third of nation Upgraded finish package (floors, countertops, millwork, fixtures) New home $372,000	Local schools ranked in middle third of nation Premium finish package (floors, countertops, millwork, fixtures) Home built in 1990 $539,000	Local schools ranked in bottom third of nation Standard finish package (floors, countertops, millwork, fixtures) Home built in 1990 $374,000
○	○	◉

The Calibration Concepts Section:

As is the practice for ACA, products can be shown to respondents who indicate a purchase likelihood score, or rate the products on a Likert scale. This section is optional within ACA, and it is optional for ACBC as well. We included a calibration concepts section in this current study, but did not use the data in estimating or calibrating part-worth utilities.

Analysis:

We estimated part-worth utilities for ACBC and CBC using our CBC/HB software for hierarchical Bayesian estimation. We employed the coding methodology as described in Johnson and Orme (2007).

The idea of including "Most Likelies" with ACBC instead of BYO is new with this paper, and it requires slightly different coding than for BYO. When coding BYO, we treat each selection for each attribute as a separate choice task. We assume the respondent is choosing among the levels only within that attribute, traded off versus price. So, there are only two columns in the design matrix that are varying: the attribute in question and price. With "Most Likelies," we code the task in the same manner, but price remains constant (at zero).

The complication with coding the "Most Likely" section is that it seems at odds with good utility theory. For example, let's assume I would prefer to purchase a home in an area with top-notch schools, but I realize I probably couldn't afford to do so. In reality, the utility of top-notch schools exceeds that for average schools, but the information contributed by the coded "Most Likely" task suggests that, *ceteris paribus*, the average school level is preferred. Thus, the information from this task could misinform the final part-worth utilities.

We investigated whether the final part-worth utilities for respondents receiving the ACBC questionnaire with the "Most Likely" section were being damaged by this problematic assumption. Three of the attributes (green/efficient construction, schools, and age of home) seemed to have *a priori* rational preference order. When we removed the three tasks associated with the most likely choices for these three attributes from the estimation, the hit rates dropped, suggesting that as-is, the tasks are quite valuable for predicting holdout tasks.

A second way to deal with this issue is to impose utility constraints on attributes with known rational utility order. We decided that attributes 2 (efficiency), 7 (schools), 9 (age of home), and 10 (price) should have rational utility order, and repeated HB analysis with constraints on those attributes. The hit rates did not improve, and in fact declined slightly (though not significantly).

So, we are left with a concern regarding the treatment of "Most Likely" tasks in the analysis. The idea of imposing constraints seems the most reasonable way to directly deal with the issue, but even ignoring the issue doesn't seem harmful for this dataset.

Holdouts:

Hit Rates

It is common for researchers to include holdout CBC tasks and to report hit rates for those tasks. As we discussed in Johnson and Orme (2007), this validation exercise stacks the deck against ACBC. Methods bias strongly favors CBC when predicting CBC-looking holdouts. Furthermore, if a key problem with standard CBC tasks is that they encourage respondents to answer lazily using aggressive simplification heuristics (that they wouldn't use in real-world purchases), then CBC utilities should have an even stronger edge in being able to predict CBC holdouts. For example, if the respondent answers each CBC task simply by choosing the home

that has 1750 square feet, then it should be quite easy to predict how this respondent would choose in holdout questions—assuming the common practice of "minimal overlap" design strategy, where only one home per task has 1750 square feet.

Because researchers expect to see comparisons between methods based on holdout predictions, we have included them in this study as well. But, there is no compelling reason to expect that ACBC should predict CBC-looking holdouts better than CBC. If CBC has weaknesses (leading to unrealistic simplification heuristics), then the CBC-looking holdouts should suffer from the same problems. Ideally, we should favor a validation based on real-world choices.

To make the holdouts more challenging, we did not design the tasks using minimal overlap or level balance. We favored levels that would be more commonly found in the typical home. And, we conducted a small pretest with a convenience sample to help fine-tune the selections so that no home was strongly dominated by another within the same choice task. And, we repeated the idea of a "customized holdout" that we employed successfully in the last two ACBC projects. We describe that directly below.

We showed five total holdout choice tasks to each respondent (prior to the ACBC or CBC sections). The first four holdout tasks showed four homes each (and no "None"). We kept track of the respondent's choices in those first four tasks, and displayed those previously selected four homes in the final (5^{th}) choice task. Thus, a respondent who always selected a home with 1750 square feet would be faced with the choice among the four winning homes, each at 1750 square feet. This would require a deeper consideration of the attribute space (beyond just the square footage attribute) to make a final selection.

In the previous two tests, ACBC did slightly worse than CBC in predicting the non-customized holdouts (but the differences were not statistically significant). But, ACBC outperformed CBC in predicting the key customized holdout on both occasions (a statistically significant result in the first test). Given the methods bias in favor of CBC, this result is remarkable—and encouraging.

The hit rates were as follows for this test, broken out by design cell:

	Cell 1 ACBC 32 cards	Cell 2 ACBC 24 cards	Cell 3 ACBC 32 cards	Cell 4 CBC 18 tasks
Holdouts 1-4	56.0	55.4	57.6	55.3
Holdout 5	40.5	48.8*	43.5	36.9*

ACBC performed better on average than CBC in predicting the first four (fixed) holdouts, and demonstrated an even greater edge in performance in predicting the customized holdout. The only statistically significant difference ($p<.05$) is marked with an asterisk. The evidence is even stronger in favor of ACBC when we add these results to the previous two tests of ACBC vs. CBC (laptop and recreational equipment studies). In those previous tests, ACBC prediction accuracy also exceeded CBC for the customized holdout task.

Although Cell 2 uses less information, it hasn't seemed to harm predictions. We should note, however, that this is a between-respondents analysis, and we cannot control for sampling error. We can conclude that all three ACBC versions seem quite viable. It seems we can cut respondent burden without substantially decreasing predictive accuracy.

In the previous ACBC study on laptop computers, in addition to HB, we also tried a method of purely individual-level analysis to compute part-worths (monotone regression). In that study, we found that monotone regression produced hit rates about 3 to 4 absolute percentage points lower than HB estimation. This demonstrated the viability of conducting purely individual-level analysis with ACBC (without sharing information across a sample of respondents to stabilize part-worths). Using the reduced-length ACBC questionnaires in the current homes study, we found monotone regression's hit rates were about 7 to 8 absolute percentage points lower than HB estimation. This suggests that if monotone regression will be used, questionnaires should be more to the length of the laptops study (see Table 3).

We also analyzed hit rates depending on the presence of the graphic of the interviewer. With the picture of the interviewer, the average hit rate (under HB) was 54.4% vs. 51.4% without her (t=2.1). It is encouraging that including this graphic and changing the way we address respondents (from "we" to "I") may lead respondents to provide more careful answers. This finding needs to be confirmed with follow-up research.

Share Predictions from Simulations

Because all respondents evaluated the same four fixed holdout tasks (each with four product concepts) prior to viewing the customized fifth holdout, we were able to summarize the shares of choice for those tasks, across all respondents. We could then use the part-worth estimates for respondents within each design cell to predict choices (by all respondents) to the first four holdout tasks. We tuned the predictions by the scale factor, to minimize the errors in prediction (reported as Mean Absolute Error, or MAE). The MAE by design cell is as follows (lower is better):

	Cell 1 ACBC 32 cards	Cell 2 ACBC 24 cards	Cell 3 ACBC 32 cards	Cell 4 CBC 18 tasks
Holdouts 1-4	3.47	3.15	2.39	5.47

All three adaptive versions of the questionnaire have lower errors of prediction than the traditional CBC method. This again is remarkable, given that methods bias would favor CBC in predicting responses to CBC-looking holdouts. We are not aware of a statistical test to compare the differences in MAE across the four cells, so we do not know if these differences are statistically significant.

Qualitative Results:

One of the key aims of the current study was to determine if the interviewing time could be reduced for ACBC while still yielding predictions at the same level or even better than CBC. In the previous two studies (laptop and recreational equipment), the ACBC survey was 2x and 3x longer than CBC, respectively. Those interview times included the optional (and unused) calibration concept section for ACBC.

This time, we shortened the ACBC surveys relative to previous studies (see Table 3). We also recorded time with and without the optional calibration section. The median times for the core conjoint tasks were as follows (omitting the calibration section[4]):

Cell 1: ACBC (BYO + 32 cards)	400 seconds
Cell 2: ACBC (BYO + 24 cards)	330 seconds
Cell 3: ACBC (No BYO + 32 cards)	346 seconds
Cell 4: CBC (18 tasks)	230 seconds

It is interesting to note that CBC respondents completed the 18 choice tasks in an average of 230/18 = 12.8 seconds each. They *must* be employing simplification strategies to answer so quickly. As researchers interested in the quality and validity of results, we should find this result concerning. The shortest ACBC questionnaires in our experiment were 330 and 346 seconds, representing about a 50% increase in time to complete the ACBC questionnaire (rather than an estimated 100% or 200% increase in time as seen in the first two tests of ACBC vs. CBC). The data also suggest (comparing cells 1 and 3) that the BYO section takes longer than the most likelies section to complete.

In absolute terms, the data show that ACBC can be conducted on a 10-attribute study with 31 total levels in an average time of 330/60 = 5.5 minutes. This certainly doesn't represent an excessive time requirement. Granted, these online panelists probably complete surveys more quickly than other sources of sample might (such as a list of customers), but the results are nonetheless encouraging. We think we can largely dismiss the concern that ACBC takes too much respondent time relative to CBC. We'd argue that respondents tend to rush through CBC interviews, and we'd prefer a method that encourages them to take a bit more time to provide more realistic data.

As we did with the previous two tests, we included some evaluation questions to assess the respondents' survey experience. There were no significant differences between the three ACBC cells, so we have collapsed their data. The results we see below almost perfectly mirror the results we saw in the laptop test: Respondents found the adaptive CBC interview to be a better overall experience, to be more realistic, less boring, and to encourage them to be a bit more careful in their responses. (As a side note, there were no statistically significant differences on

[4] An analysis of the total time including the calibration tasks suggests that the calibration tasks take about 10 seconds on average to complete. Thus, a 5-task calibration section would be expected to add nearly a minute to the interview time for ACBC.

these qualitative variables depending on whether the graphic of the attractive interviewer was shown or not.)

How would you compare your overall experience with this survey compared to other internet surveys you have completed?

	ACBC	CBC
This survey was far better (5):	**27%**	19%
This survey was better (4):	**46%**	39%
This survey was about the same (3):	26%	**38%**
This survey was worse (2):	2%	**4%**
This survey was FAR worse (1):	0%	0%
Means:	4.0	3.7 (t = 4.8)

How much do you agree with the following statements about this survey?
(5=Strongly Agree, 1=Strongly Disagree; Top Box % shown beneath means.)

	ACBC	CBC	
Q1. The homes I was asked to evaluate seemed realistic.	**4.1** **40%**	3.9 26%	(t=4.2)
Q2. This survey was at times monotonous and boring.	**2.5** **4%**	2.8 5%	(t=3.3)
Q3. I'd be very interested in taking another survey just like this in the future.	4.3 54%	4.2 51%	(t=1.1)
Q4. The survey format made it easy for me to give realistic answers that reflect exactly what I'd do if buying a real home.	**4.2** **46%**	4.0 38%	(t=3.3)
Q5. The way the homes were presented made me want to slow down and make careful choices.	**4.0** **33%**	3.8 26%	(t=2.8)

Simplification and CBC:

This time, we decided to ask CBC respondents regarding the strategies they employed to answer the CBC tasks. First, we asked an open-end question (see Appendix C). Then, we asked two closed-ended questions, as follows:

Which of the following statements best reflects how you made choices among the houses in this survey?

76% There are certain minimum requirements I need in a house, and I used those requirements to narrow down my choices and make a decision.

18% I didn't have any absolute requirements. Even if a house wasn't ideal in some ways, if it had enough other good features, I could choose it.

5% Neither of these

Different people pay attention to different features of the houses when responding to this questionnaire. In reality, some of the features we asked you to consider might not have mattered to you. Which of the following best describes how you approached the challenge of deciding among 4 houses each described on 10 aspects/features?

14% I focused on every feature before making a decision
29% I focused on most of the features before making a decision
22% I focused on about half of the features before making a decision
23% I focused on at most 3 or 4 of the features before making a decision
10% I focused on just 1 or 2 features to make a decision
2% None of the above

The first question confirms what we would suspect for home purchases: respondents have certain requirements, and they do not behave in a purely compensatory manner. Situations like these would especially favor the ACBC methodology.

The second question also confirms the obvious: most respondents do not pay attention to all features prior to making a selection in CBC tasks. Over half of the respondents report that they pay attention to half or fewer of the features prior to making a choice. There is probably some social desirability bias in these responses. We suspect even more simplification than is reported here. Appendix C includes open-end responses from some CBC respondents, indicating simplification and non-compensatory strategies. Other researchers have shown, analytically, that a great deal of simplification appears to take place when respondents answer complex CBC tasks (Gilbride and Allenby 2004). ACBC would appear to have an advantage when we suspect respondents might be inclined to pay less attention to all the features in a market research survey than they would if making a high-consequence, real world purchase.

Conclusions:

This test confirms findings of two previous ACBC studies. We also have shown that ACBC questionnaires can be shortened significantly, without significant loss in performance (predicting CBC-looking holdouts) relative to CBC. It is possible that our improvement in the design generation algorithm is partly responsible, though we did not include a control group to verify this. It is also possible that home purchasing decisions, being quite subject to non-compensatory cutoff rules, tipped the scales more in favor of ACBC (over CBC) than for the previous studies involving laptops and recreational equipment.

The (shortened) ACBC questionnaires in this study were about 50% longer than the CBC interview (rather than 100% or 200% longer as in the previous two studies). Despite the relative decrease in questionnaire length, predictive validity exceeded CBC (for CBC-looking holdouts, which would naturally favor the CBC model). Furthermore, respondents found ACBC more engaging, realistic, and a better overall interviewing experience, than standard CBC.

One should not conclude from this research that ACBC questionnaires can be shortened without cost. Although predictions of the CBC-looking holdout tasks did not seem to improve when using longer questionnaires, we believe the responses to our holdouts reflect unrealistic simplification strategies that do not take advantage of the added depth of information that ACBC can provide. ACBC does better than CBC in quickly learning which levels are must-have features, taking that as given, and then probing deeper regarding the more compensatory tradeoffs for attributes of next importance. If the goal of the research is to obtain very accurate classification/prediction of individual respondents in terms of more realistic real-world decisions, we would prefer ACBC questionnaires that went beyond the abbreviated exercises investigated here with their minimal 5- to 6-minute requirement. ACBC questionnaires are more engaging than CBC, and if we can afford a few extra minutes to learn more in depth about respondents' preferences, a longer interview more in line with our previous two ACBC tests would be palatable to the respondent and valuable for the researcher.

We plan to release a beta version of ACBC later this year, and hope this will open the door for more people to test this promising methodology. We also hope that future tests might involve validation to real-world purchases. Under these conditions, it is likely that ACBC will continue to demonstrate an edge over CBC.

Appendix A:
CBC Tasks Layout

Assuming these were the only available options, please tell me which of the following homes you would likely purchase?

(3 of 18)

1400 Square Feet	1750 Square Feet	3000 Square Feet	4000 Square Feet	
Built green certification & superior efficiency	Standard construction & efficiency	Built green certification & superior efficiency	Standard construction & efficiency	
2 bedrooms, 1.5 bath	2 bedrooms, 1.5 bath	4 bedrooms, 3.5 baths	3 bedrooms, 2.5 baths	
Medium master bath (separate shower and tub), one walk-in closet	Small master bath (integrated shower/tub), small walk-in closet	Small master bath (integrated shower/tub), small walk-in closet	Large master bath (separate shower & deep soaking tub), two walk-in closets	
2-car garage	3-car garage	3-car garage	2-car garage	None of these homes comes very close to fitting my needs and budget.
Large lot (0.5 acres)	Small lot (0.2 acres)	Small lot (0.2 acres)	Medium lot (0.35 acres)	
Local schools ranked in bottom third of nation	Local schools ranked in top third of nation	Local schools ranked in top third of nation	Local schools ranked in middle third of nation	
Upgraded finish package (floors, countertops, millwork, fixtures)	Standard finish package (floors, countertops, millwork, fixtures)	Standard finish package (floors, countertops, millwork, fixtures)	Premium finish package (floors, countertops, millwork, fixtures)	
Home built in 2000	New home	Home built in 1990	Home built in 1980	
$236,250	$398,500	$379,750	$715,000	
○	○	○	○	○

Appendix B
Counts: BYO vs. "Most Likelies"

	BYO	Most Likelies
1400 sq ft	24	11
1750 sq ft	35	32
2250 sq ft	28	41
3000 sq ft	9	9
4000 sq ft	4	6
Std cons.	65	60
Green cons.	35	40
2bed, 1.5bth	13	10
3bed, 2.5bth	49	53
3bed, 3bth	19	19
4bed, 3.5bth	14	15
5bed, 3.5bth	4	3
Sml bath	18	19
Med bath	58	54
Lrg bath	24	27
2-car garage	86	83
3-car garage	14	17
Sml lot	17	15
Med lot	58	59
Lrg lot	24	27
Schools bottom	12	4
Schools mid	61	56
Schools top	27	40
Std finish	47	38
Upgrade finish	41	47
Premium finish	13	15
Built in 1980	15	12
Built in 1990	42	34
Built in 2000	25	30
New home	18	24

Notes: This table shows the percent of times respondents selected each level for each attribute as the characteristic they would like to have in a home they wished to purchase. With the BYO respondents, prices were associated with each level, as shown earlier in Table 2. With respondents who received the "Most Likelies" section (instead of BYO), no prices were attached to the attribute levels. While there are statistically significant differences between the columns (as would be expected, since price penalties were not directly shown with upgraded levels for "Most Likelies"), the overall patterns of preferences for homes were quite similar. Respondents to the Most Likelies seem to have been able to anticipate reasonable costs for the upgraded features, and choose a realistic level based on the combination of their budget and preferences.

Appendix C
Respondents Describe How They Answered Standard CBC Questions

We asked respondents who completed the standard CBC tasks what strategies they used to answer the 18 choice tasks, each with 4 alternatives plus "None". Most of the responses indicate non-compensatory or simplification heuristics at odds with CBC's additive logit rule assumption. Here are some examples of non-compensatory strategies that were expressed (edited to correct some grammatical errors and standardize the capitalization):

What I did was I eliminate houses that were definitely not me such as price range and very low square feet first. By doing that, I am not considering a home that I would never buy anyway, so then I can consider a home that I really want.

I looked at the price, size and other main factors that were very important and used them to "weed out" some of the options. Then I read through the remaining options and picked the best one.

I focused on the most important things for me. The number of rooms was the most important so if the number was unacceptable then I took it out of my consideration, then I looked at the yard size and so on and so on.

There are certain aspects that I find most important when choosing a home so I looked at the row for those characteristics first. If one of the columns had something I really did not like in that important row, I immediately took it off the list.

I would first rule out by price, nothing over $300,000. Then I looked at the schooling and dropped any that were in the bottom third of the nation. After that I would look at the lot size and home size to make my final decision.

I knew which things I would rule out right away--the 1400, 3000, and 4000 sq ft homes; any homes with less than 2 baths; any in school districts in bottom third; any over $300,000. Then I could look at whichever houses were left and see if they had other qualities I liked.

I looked at the categories that were most important to me. For example, I wanted a large lot and a new home--those two categories were first looked at. If the match wasn't found, then I'd expand to other categories that were important and keep sorting the information until I made my choice.

I looked at the different options and eliminated them according to how I felt. I didn't want too small of a house so I never went for ones that were below 2,000 sq ft. So, I usually only had two options that I looked over more carefully.

I started by looking at houses that fit my budget and then narrowed it down from there.

First I looked at price. I ruled out the houses that were priced too high for my budget, so I didn't have to read their information. Then I looked at size of the remaining houses. If it was too small, I ruled that out and didn't have to read the rest of the information. Usually then there was

only one house left. If it fit my needs I would select that as my choice. If not, I'd say none of the houses fit my needs.

First is to look at the prices. If the price is out of my budget, it can easily eliminate one or two of the houses, depending on if the price is too low or too high. Then, I look at how much square footage I need, if there is less than 2000 sq. ft. I probably won't buy, but I don't need too much more than that. Next, I look at the features of the house (year built, quality, green-ness) and the schools to make my choice. When going through it in this systematic way, it makes the choices quick and easy.

I automatically eliminated houses that were not new & were under 1750 square feet.

I checked the price and size first, and if those were in my range then I would read the rest of the information. If the price and or size did not fit then I did not bother because the rest would not have mattered.

First I looked at the prices for all of the houses, then after I found the ones that were in my price range, I scrolled my eyes upward to look at the other features, and from there I chose which one I would buy.

First I zoomed in the all school choices - ranked top, middle, or bottom. I eliminated all houses that fed into a bottom third school. Next I looked at the type of finishes and the year the house was built. Lastly, I compared price to see if there was a best fit.

I first picked the type of home I wanted, (green) and then went from there. That eliminated 50%.

I looked at the must have features first and eliminated the houses that didn't have them.

I picked the few things that I could not live without. There was usually only one option that had them all.

I eliminate ones first based on how many bedrooms I would eliminate them if there were less than three, then I jump to the school--I eliminated the ones that were in the bottom third of nation.

I chose to read each line separately. First, I scrolled all the way to the bottom to look at the price. I found the few selections that were in my budget, then looked at their school districts. If it was a bad district I would not consider it. Next, I looked through the other descriptions, and from there I would decide which house I liked best based on square footage, bedrooms, and price.

References:

Gaskin, S., Evgeniou, T., Bailiff, D. and J. Hauser (2007), "Two-Stage Models: Indentifying Non-Compensatory Heuristics for the Consideration Set then Adaptive Polyhedral Methods within the Consideration Set," Sawtooth Software Conference Proceedings, Sequim, WA.

Gilbride, Timothy and Greg M. Allenby (2004), "A Choice Model with Conjunctive, Disjunctive, and Compensatory Screening Rules," *Marketing Science*, 23, 3, (Summer) 391-406.

Johnson, Rich and Bryan Orme (2007), "A New Approach to Adaptive CBC," Sawtooth Software Conference Proceedings, Sequim, WA. Also available at www.sawtoothsoftware.com/techpap.shtml.