Do homes with better energy efficiency ratings have higher house prices?

Experimental approach
Preface

Copenhagen Economics was commissioned by The Danish Energy Agency to examine the relationship between house prices and energy standards. The analysis comprises both an econometric and an experimental approach. The final result is presented in three separate reports: One on results and methods from the econometric approach, another on results and methods from the experimental approach and a third summarizing the key results from both approaches. You are now reading the report on results and methods from the experimental approach.

This background report explains in detail the methods we applied within the experimental part and presents our results. We conducted three different types of experiments to assess the effect of the energy standard on the house price. Methods and results from these three experiments are explained in the following three chapters, respectively the web-based experiment, the physical experiment and the conjoint analysis.
# Table of contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preface</td>
<td>1</td>
</tr>
<tr>
<td>1  Previous findings and expectations</td>
<td>5</td>
</tr>
<tr>
<td>2  The web-based experiment</td>
<td>7</td>
</tr>
<tr>
<td>2.1 Why a web-based experiment?</td>
<td>7</td>
</tr>
<tr>
<td>2.2 Setting and design of the experiment</td>
<td>7</td>
</tr>
<tr>
<td>2.3 Results</td>
<td>11</td>
</tr>
<tr>
<td>3  The physical experiment</td>
<td>18</td>
</tr>
<tr>
<td>3.1 Why a physical experiment?</td>
<td>18</td>
</tr>
<tr>
<td>3.2 Setting and design of the experiment</td>
<td>18</td>
</tr>
<tr>
<td>3.3 Method / Testing for an effect</td>
<td>21</td>
</tr>
<tr>
<td>3.4 Results</td>
<td>22</td>
</tr>
<tr>
<td>4  The conjoint analysis</td>
<td>31</td>
</tr>
<tr>
<td>4.1 Why a conjoint analysis?</td>
<td>31</td>
</tr>
<tr>
<td>4.2 Idea in combination with the other experiments</td>
<td>32</td>
</tr>
<tr>
<td>4.3 Setting and design of the survey</td>
<td>33</td>
</tr>
<tr>
<td>4.4 The model</td>
<td>40</td>
</tr>
<tr>
<td>4.5 Results</td>
<td>42</td>
</tr>
<tr>
<td>References</td>
<td>49</td>
</tr>
</tbody>
</table>
List of tables

Table 1 Relationship and effects in previous literature .................. 5
Table 2 The 24 combinations across energy label and price ............ 10
Table 3 The 10 combinations across components, knowledge and roof ................................................................. 10
Table 4 Information provided for each house ............................. 19
Table 5 Energy labels and house values .................................... 21
Table 6 Variances in the house price assessments ....................... 30
Table 7 Attributes and their levels .......................................... 34
Table 8 Details on the explanatory variables ............................. 39
Table 9 Logit regression results ............................................. 43
Table 10 Marginal changes in odds ........................................ 45
List of figures

Figure 1 Theoretical price increase related to energy labels ..............6
Figure 2 Example of the house particulars ........................................... 8
Figure 3 Relation between house price and energy standard ................. 11
Figure 4 Relation between house price and energy standard assessed by real estate agents ............................................................... 12
Figure 5 Relation between house price and energy standard assessed by Danish citizens ................................................................. 13
Figure 6 Effect of monetary information about energy labels ............... 13
Figure 7 Effect of monetary information for Danes and agents ............ 14
Figure 8 Effect across house prices (label change F → C) .................. 15
Figure 9 Effect of new windows and cavity wall insulation ............... 15
Figure 10 Effect of high energy standard vs. roof replacement .......... 16
Figure 11 Geographical differences in effect (label F → C) ............... 17
Figure 12 Geographical differences across Danes and agents .......... 17
Figure 13 Aggregated results for raw and cleansed data .................... 23
Figure 14 Comparison of average house prices ................................... 24
Figure 15 The energy label effect for raw data .................................. 25
Figure 16 The variance of the house price assessments ....................... 26
Figure 17 The energy label effect for cleansed data ............................. 28
Figure 18: The energy label effect for low-variance-houses.............. 29
Figure 19 The four different choice modelling methods ..................... 32
Figure 20 The perfect size of the fractional factorial design ................. 35
Figure 21 Differently ordered versions of the same alternative......... 40
Figure 22 The cumulative distribution function in a logit model .......... 41
Figure 23 The trade-off between energy label and house size .......... 43
Figure 24 The effect of the energy label on the choice ....................... 46
Chapter 1
Previous findings and expectations

In the past decade the relationship between energy standards and house prices has been widely discussed. Contributions are many but primarily stem from the real estate business, mortgage credit institutions and relevant public sector departments. A concrete outcome is a number of analyses that — despite their different approaches — all find a positive relationship between energy standards and house prices, cf. Table 1. For example do the econometric analyses estimate a D-labelled house to have a price 5-10 per cent lower than an A-labelled house and 5-19 per cent higher than a G-labelled house.

Table 1 Relationship and effects in previous literature

<table>
<thead>
<tr>
<th>Approach</th>
<th>Relationship (+/-)</th>
<th>Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Econometric</td>
<td>+</td>
<td>Sales price relative to a D-label: A/B (5 %), C (1,8 %), E (-0,7 %)</td>
</tr>
<tr>
<td></td>
<td>+</td>
<td>Sales price relative to a D-label: A/B (6,4 %), C (6 %), E (-6,2 %), F (-12,3 %)</td>
</tr>
<tr>
<td></td>
<td>+</td>
<td>3 times higher price rebate for a G house compared to C house</td>
</tr>
<tr>
<td></td>
<td>+</td>
<td>Higher square meter prices and lower price rebate for efficient houses</td>
</tr>
<tr>
<td>Qualitative</td>
<td>(+)</td>
<td>&quot;Limited impact from energy on house market prices&quot;</td>
</tr>
<tr>
<td></td>
<td>(+)</td>
<td>&quot;Overall it is an open question how much the price effect will eventually be - other than it is hardly controversial to conclude that there will be a positive effect, but it will be much less than the theoretical effect&quot;</td>
</tr>
</tbody>
</table>

Note: Effects from econometric analyses were all significant

The discrepancies in terms of reflect the difficulties of credibly assessing the effect from energy standards on house prices. In fact, the vast majority of previous studies limit themselves to investigate only the relationship and not the causal relationship between energy standards and house prices. So whether an increase in energy standards in fact causes an increase in the house price is still of limited evidence.

Acknowledging the difficulties of quantitatively analysing the question of interest the acceptance for sheer qualitative analyses rises. In this study we therefore started out by conducting a round of interviews thereby adding to the number of qualitative analyses in the table above. We asked relevant parties of the Danish real estate business to give their opinion on the relationship between energy standards and house prices. Apart from gathering valuable information shaping expectations this also served the purpose of quality assuring our planned experiments.

The general assessment by the real estate business is very much in line with previous qualitative findings. The real estate agents believe there is a positive correlation between energy
standards and house prices. Especially among owners of estate agencies and representatives of The Danish Association of Chartered Estate Agents (Dansk Ejendomsmæglerforening) the opinion was that higher energy classes provide a faster sale and a smaller rebate from the initial asking price. In particular, houses with the lowest energy labels (G, F, and partly D) were found hard to sell.

Our initial interviews also reveal that real estate agents believe that the positive correlation is more pronounced in cities, as people are more environmentally conscious and generally higher educated (and thus more aware of energy costs). The agents believe that visible energy renovations, such as new A-labelled windows, can be expected to have a larger importance than invisible renovations, such as cavity wall insulation.

In addition to the previous findings and our own qualitative analysis, the theoretical relationship between energy standards and house prices is highly relevant when forming expectations to the coming results of this study. The theoretical effect from increasing the energy standard by one energy label is 71,000 DKK on average with decreasing effects as the label increases (ranging from 41,000 to 91,000 DKK), cf. Figure 1.

**Figure 1 Theoretical price increase related to energy labels**

<table>
<thead>
<tr>
<th>Energy Label</th>
<th>Extra Price Compared to a Similar G-labelled House</th>
</tr>
</thead>
<tbody>
<tr>
<td>F</td>
<td>91,000</td>
</tr>
<tr>
<td>E</td>
<td>81,000</td>
</tr>
<tr>
<td>D</td>
<td>172,000</td>
</tr>
<tr>
<td>C</td>
<td>247,000</td>
</tr>
<tr>
<td>B</td>
<td>326,000</td>
</tr>
<tr>
<td>A</td>
<td>385,000</td>
</tr>
</tbody>
</table>

Note: Energy label G is used as reference point. We use average energy prices of 0.69 DKK per kWh, considering a house of 100 sq. m. and each energy label refers to the average energy consumption within the threshold. We account for the average values of options to renovate.

Source: Copenhagen Economics (2015) Danish house prices and the effects of energy standards – econometric approach

Based on the previous literature, the opinions of the real estate business and the theoretical effects we expect in this study to find a positive effect from energy standards on house prices. We expect the effect to be largest among the lowest energy labels both due to the theoretical savings and on the basis of the ‘kink’ argument from the real estate business. Finally, we expect to find a larger effect in cities and that visible renovations have larger importance than invisible renovations.
Chapter 2
The web-based experiment

2.1 Why a web-based experiment?
The web-based experiment was designed to test for effects of energy standards on house prices in a setting mimicking a standard web page for house sales. As the only experiment, this was not only carried for real estate agents but also for Danish citizens. By varying the energy standard in otherwise identical house sales particulars, we could causally assess the quantitative effect of energy standards on house prices.

In essence, this experiment therefore answers the question: How large a value do real estate agents and average Danes attach to energy standards? The question is answered by comparing the assessed house price of a house with, say, energy label C to the exact same house but with energy label D.

One important strength of this experiment is its proximity to real-life estate web pages where potential buyers based on few house characteristics decide whether to click on a house offer and declare their interest. The experiments’ main strength is its setting: With 17,000 observations from more than 1,500 individuals in a highly controlled environment due to the randomized web approach the certainty of detecting even smaller causal effects is quite high.

2.2 Setting and design of the experiment
This section contains a description of the practical execution (setting) followed by an explanation of the design of the experiment.

We conducted the web-based experiment in cooperation with Dansk Ejendomsmæglerforening (The Danish Association of Chartered Estate Agents) and Kompas Kommunikation. Dansk ejendomsmaegetorenforening helped us by sending out invitations for participating in the web-based questionnaire to 2,328 real estate agents in Denmark. Kompas Kommunikation helped us by sending out invitations to Danish citizens obtaining a geographically representative sample for Denmark. Responding the questionnaire was incentivized by promising 10 gift cards each valuing 500 DKK to be distributed at random.

By accepting the invitation the respondents click on a link that directs them to a web page containing the housing questionnaire. Here the respondents are met by an introduction to the questionnaire followed by eight or nine consecutive particulars for which they each time are asked to guess the house price.

The introduction explicitly states that the questionnaire is a crucial part of research project and that it is important the respondents give well-considered answers. The respondents are asked to answer seven simple background questions addressing gender, age, education, postcode, current housing type, income and self-estimated housing market knowledge. In the final part of the introduction, the respondents are informed that they will be shown a
number of particulars for which they in each case have to guess the value the specific house has just been sold for. Finally, they have the option to state their e-mail address in order to participate in the lottery of winning 500 DKK.

Each of the particulars contain key information replicating the standard setup for actual web page housing particulars, cf. Figure 2. The setup for all particulars is identical but contains different housing information.

**Figure 2 Example of the house particulars**

**Boligopslag 1**

![Image of house particulars](image_url)

**Fakta om boligen**
- Type: Villa
- Beliggenhed: Sorgenfri
- Boligareal: 135 m²
- Grundareal: 448 m²
- Rum: 4
- Antal plan: 2
- Byggeår: 1926
- Energimærke: ☺

Pæn entré, hyggeligt køkken/alrum. To meget flotte lyse stuer med flotte plankegulve. 1. salen indeholder stor repos med plads til kontor, stort lyst værelse samt stort lekkert soveværelse.

Sidste renovering blev foretaget i 2008.

Source: Copenhagen Economics

All particulars contain a picture, key facts and a text describing the house. Among the facts are listed location, square meters of living area, construction year, etc. The text allows for specific remarks about the house, e.g. the impending need for a roof replacement.

In the case of citizens, respondents are informed that the house was set for sale at a certain price, e.g. 3 million DKK, and has just been sold. The respondents are then asked how much they think it was sold for, and how much they would be willing to pay themselves. In the case of real estate agents, respondents are informed that similar houses in the same area have been valued at a certain price, e.g. 3 million DKK, by a local real estate agent. They are then asked how much the think the house is sold for. Thus, this experiment addresses the link between energy standards and the actual house price not to be confused with the asking price.

In order to test several relevant hypotheses and not have too many questions in just one questionnaire, we constructed four questionnaire varieties. Two consisted of eight particulars and the other two of nine. This gave in total 34 unique particulars or houses. Each
respondent were asked for the price for either eight or nine of these particulars. The particulars differed from each other by three dimensions: Energy standard, price category and remarks.

**1. Energy standard.** We use four different energy labels to reflect the energy standard: C, D, E and F. By changing the energy label we can measure the effect from increasing the energy standard, e.g. from D to C. We can also measure the effect across several label levels, e.g. from F to C. The labels A, B and G are not included since relatively few houses have these labels and virtually none exists of newer date.

**2. Price category.** We use three price categories: 2, 3 and 4 million DKK. The purpose of having different price categories is to test the effect of energy standards across house prices. Thereby we are able to answer whether the absolute effect from jumping from D to C is greater for a 4 million DKK house compared to a 2 million house. This is interesting since the energy standard in theory should have the same absolute effect irrespective of house prices all other things being equal. Only if the more expensive house has more square meters the effect should be greater. However, this was not the case in this experiment. Results of this test will allow to conclude whether the absolute effect is in fact independent of the house prices corresponding to that the relative effect is decreasing in house prices.

Since the respondent has to guess the house price, we cannot state the price category explicitly. Instead, we indicate the price by changing the quality of the house via the picture of the house, the text and the facts about the house. The text is changed by highlighting different features such as a completely new kitchen and the facts are changed by for instance increasing the number of square meters. See appendix 1 for an overview of house pictures and texts used. Finally, the respondent is told the asking price which differs between 2, 3 and 4 million DKK. This indicates the price level as well.

**3. Remarks.** We use five different types of remarks, listed in Danish as they were presented:

1. ‘Sidste renovering blev foretaget i 2006.’
2. ‘Sidste renovering blev foretaget i 2006. Huset blev her udstyret med A-mærkede energivinduer.’ We label this remark ‘windows’.
3. ‘Sidste renovering blev foretaget 2007. Huset blev i den forbindelse hulmursisoleret.’ We label this remark ‘cavity wall’.
4. ‘Sidste renovering blev foretaget i 2010. Det forventes dog, at taget skal renoveres indenfor de næste 5 år.’ We label this remark ‘roof’.
5. ‘Sidste renovering blev foretaget i 2008. I forbindelse med renoveringen blev der foretaget energiforbedringer, der førte til, at energimærket steg fra F til E. Dette har givet en samlet energibesparelse på mere end 5.000 kr. om året sammenlignet med tidligere.’ We label this remark ‘knowledge’.

The first remark is a standard remark. The year varies slightly but the variation is not used for analysis. Remark 2 and 3 about windows and cavity wall insulation (hulmursisolering) allow us to test the importance of which components yield the energy standard. In other words, has a C-house with new windows a higher value than a C-house with cavity wall insulation? That we can determine by analysing the effects of these remarks.
Remark 4 concerning the roof replacement has the purpose of testing the importance of energy standards when a different crucial factor enters the picture. It might be that there is an effect of the energy standards that disappears when an important depreciating factor is introduced.

Remark 5 concerning the monetary consequence of the energy standard has the purpose of increasing knowledge about the importance of different energy labels. If additional information increases the house value we can conclude that lacking knowledge or attention is limiting the effect of energy standards.

For each of three price categories we used two pictures of the same house, labelled A and B. By presenting the exact same house from the front and back we could hide the fact that we only presented three different houses in total. Four energy labels (C, D, E and F), three price categories (2, 3 and 4 million DKK) and two picture types (A and B) yields in total 24 combinations of unique house types or particulars. These were then distributed across the four questionnaires, cf. Table 2.

<table>
<thead>
<tr>
<th>Energy label / Price</th>
<th>2,A</th>
<th>2,B</th>
<th>3,A</th>
<th>3,B</th>
<th>4,A</th>
<th>4,B</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>D</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>E</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>F</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>1</td>
</tr>
</tbody>
</table>

Note: The numbers in the cells refer to in which of the four questionnaires the combination is included. A and B refer to which of the two pictures of the same house that is used. For example the 4 million house with energy label E and photo type A is present in questionnaire 1.

Source: Copenhagen Economics

To be able to test the effect from different remarks we constructed yet another 10 combinations, cf. Table 3.

<table>
<thead>
<tr>
<th>Energy label / Price</th>
<th>2,A</th>
<th>2,B</th>
<th>3,A</th>
<th>3,B</th>
<th>4,A</th>
<th>4,B</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: The terms in the cells refer to which remark is provided in the given combination. For example the remark about roof is present in the combination of a 2 million house with energy label C using picture type A. The numbers in brackets refer to in which of the four questionnaires the combination is included. Using the same example this house would be in questionnaire 3.

Source: Copenhagen Economics
All forthcoming figures are based on the results from the above two tables. The method is quite straightforward: The average effect of improving energy standards from, say, D to C is simply the difference of the averages of the price corresponding to the first two rows of table 2. When analysing the effect across house price categories (2, 3 and 4 million DKK), averages are calculated separately for the three groups in table 2, that is 2,A and 2,B together versus 3,A and 3,B together versus 4,A and 4,B together. Table 3 is used when analysing the effects of remarks. As can be seen remarks are only present for some energy labels and some house price categories. In the case of ‘knowledge’ for example, we compare the average of a standard E-labelled 4 million house with the average of an E-labelled 4 million house with knowledge.

For the analyses we applied a one-tailed, two-sample t-test to test whether the difference in energy standard had a significant effect on the house price. The same test was used when applied to testing the effects of remarks and geographical aspects. The latter refer to testing whether respondents from different parts of Denmark value energy standards differently. This is the only analysis that uses more information than presented in table 1 and 2, namely the answers from the background question on post code. This does not give rise to adding new particulars. When testing knowledge, that is the effect of monetary information, we applied a paired t-test instead of unpaired since it was the same individuals who were presented for the two relevant houses.

### 2.3 Results

The web-based experiment indicates that energy standards have a positive, causal and significant effect on house prices, cf. Figure 3.

**Figure 3 Relation between house price and energy standard**

- **DKK.**
  - 2,800,000
  - 2,850,000
  - 2,900,000
  - 2,950,000

- **Price Differences:**
  - F to E: 8,000
  - E to D: 53,000
  - D to C: 41,000

**Note:** Results are for both estate agents and Danes. Significance is evaluated in relation to the energy label below.

**Source:** Copenhagen Economics
Real estate agents and average Danish citizens estimate that house prices are higher when the energy label is better. The house price increases stepwise as the energy label increases from F to C. The largest increase is from E to D (53,000 DKK) and from D to C (41,000 DKK) both being highly significant. First step from F to E measures 8,200 DKK and is not significant. In total, the difference between lowest and highest energy standard (F and C) is 101,000 DKK.

Compared to the theoretical expectations presented in chapter 1, the found effects are considerably lower. For the jumps from E to D and D to C the effects are 29 per cent and 48 lower than expected, respectively. The jump from F to E is insignificant which is surprising given that we expected this jump to yield the largest increase.

Looking only at real estate agents the same step-wise pattern emerge, cf. Figure 4. The main difference is that the real estate agents assess the effect from jumping from F to E as larger (25,000 compared to 8,000). Also, the house prices are in general estimated somewhat higher. This reflects that the agents make their own independent judgement and are not influenced by the stated asking price as much as Danes in general.

**Figure 4 Relation between house price and energy standard assessed by real estate agents**

<table>
<thead>
<tr>
<th>DKK</th>
<th>2.900,000</th>
<th>3.000,000</th>
<th>3.100,000</th>
<th>3.200,000</th>
</tr>
</thead>
<tbody>
<tr>
<td>F</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>25,000**</td>
<td>63,000***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>39,000***</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Results are only for real estate agents. Significance is evaluated in relation to the energy label below. Source: Copenhagen Economics

Looking only at Danish citizens we see the same step-wise effects except for an insignificant jump from F to E, cf. Figure 5. It is important to acknowledge what insignificance implies, namely that what may to some seem as a negative effect of 8,000 DKK is in fact an effect of zero – the value could just as well had been a positive insignificant 8,000 DKK. One explanation is that energy standards matter less when improving among the lowest energy standards. In other words there is a threshold effect that creates a kink when improving beyond E. As was the case for the real estate agents both the jumps E-D and D-C are reasonably large and significant.
Monetary knowledge about energy standards has great impact on house prices. When estate agents and Danes are explicitly informed about the monetary effect of energy standards the estimated price is 82,000 DKK higher, cf. Figure.

Figure 5 Relation between house price and energy standard assessed by Danish citizens

<table>
<thead>
<tr>
<th>DKK</th>
<th>2.650.000</th>
<th>2.700.000</th>
<th>2.750.000</th>
<th>2.800.000</th>
</tr>
</thead>
<tbody>
<tr>
<td>F</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>8,000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>42,000***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>43,000***</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Results are only for real estate agents. Significance is evaluated in relation to the energy label below.
Source: Copenhagen Economics

Figure 6 Effect of monetary information about energy labels

<table>
<thead>
<tr>
<th>Kr.</th>
<th>3.700.000</th>
<th>3.750.000</th>
<th>3.800.000</th>
<th>3.850.000</th>
<th>3.900.000</th>
</tr>
</thead>
<tbody>
<tr>
<td>without knowledge</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>with knowledge</td>
<td>82,000***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: The effect is measured for a 4 million DKK house with energy label E. The text informs that the energy label has risen from F to E during the last renovation which causes a yearly savings of 5,000 DKK.
Source: Copenhagen Economics
The effect of monetary information about the energy label is highly significant for both real estate agents and Danish citizens and, perhaps a bit surprising, greatest for the real estate agents, cf. Figure 7.

**Figure 7 Effect of monetary information for Danes and agents**

Note: The effect is measured for a 4 million DKK house with energy label E. The text informs that the energy label has risen from F to E during the last renovation which causes a yearly savings of 5,000 DKK.

Source: Copenhagen Economics

Ideally, the energy label should reveal all relevant information about the current energy standard of the house, and information on historical energy improvement should theory be irrelevant. However, the fact that “knowledge” affects prices (figure 6 and 7), could reflect that agents are not fully informed with the energy label.

In monetary terms the energy standard has the greatest effect on more expensive houses (houses priced at 3 and 4 million DKK) compared to less expensive houses (houses priced at 2 million DKK), cf. Figure 8. Same conclusion remains when looking at square meter effects (not shown in the figure). The difference disappears when considering the effect relative to the house price – here only the 3 million house has a significant positive effect.
Explicit information about new cavity wall insulation has great effect, while new A-labelled windows do not increase the house price significantly, cf. Figure 9.
The distinction between cavity wall insulation and windows was designed to test the hypothesis that visible energy renovations like windows are more important than invisible renovations like cavity wall insulation. We find the opposite. One possible explanation is that visibility plays a role in real life, where the estate agents in a short time themselves determine the energy standard. However, in this experiment everything is visible; cavity wall insulation is explicitly stated. Though cavity wall insulation matters the most in the experiment it is therefore not certain that this is the case in reality since the experiment unfortunately does not mimic real life very well in this test. In our opinion we therefore cannot determine whether new A labelled windows or cavity wall insulation matter the most in real life. It seems clear though, that explicit information about energy renovations, such as cavity wall insulation, can have great impact on whether it is accounted for in the price assessment or not.

When the need for roof replacement is stated the house price drops significantly, cf. Figure 10. Information about the roof makes the price fall by 96,000 DKK, which might reflect the future expected cost for a new roof. This drop for a C labelled house is greater than the difference between a C and an F labelled house. This indicates that unpleasant or alarming information may cause the effect of energy standards to become negligible.

**Figure 10 Effect of high energy standard vs. roof replacement**

![Figure 10 Effect of high energy standard vs. roof replacement](image)

Note: A 2 million DKK house is applied for the analysis of the roof effect. Significance is for label C measured compared to label F and for label C, bad roof compared to label C

Source: Copenhagen Economics

Region Hovedstaden is the Danish region that attributes the lowest value to energy standards, cf. Figure 11. The effect of jumping from energy label F to C is valued 2-3 times lower in Region Hovedstaden compared to the rest of the country.
The geographical differences are more or less the same for both real estate agents and other Danes, cf. Figure 12. For both groups Region Hovedstaden is clearly the Region that treasures energy standards the least.

Figure 11 Geographical differences in effect (label F → C)

Note: The price difference between an F and a C house is compared across different regions in Denmark.
Source: Copenhagen Economics

Figure 12 Geographical differences across Danes and agents

Note: The price difference between an F and a C house is compared across different regions in Denmark.
Source: Copenhagen Economics
Chapter 3
The physical experiment

3.1 Why a physical experiment?
The physical experiment was designed to test for an effect of the energy label on the house price in a real world setting. Two randomised groups of real estate agents assess the value of the same houses, the only difference being the energy label, which was manipulated up- or downwards for one of the groups. If the energy label affects the house price, then we expect to see a higher assigned value (on average) in the group that was presented the better label.

Mimicking a real-life situation, the physical experiment supplements our other more theoretical experiments. It provides a valuable insight into if (and if yes, how) real estate agents consider the energy label when they assess the house price in reality. We observed the real estate agents’ behaviour for real houses without simplifying the situation. The experiment’s proximity to reality is its main strength.

3.2 Setting and design of the experiment
This chapter contains a description of the practical execution (setting) followed by an explanation of the design of the experiment.

We conducted the physical experiment in cooperation with the Dansk Ejendomsmaælgerforening (The Danish Association of Chartered Estate Agents) as part of their workshop on the topic “Boligens udbudspris”. The workshop took place in Vejen and Roskilde and aimed at improving the participants’ skills in assessing house prices correctly by appropriately considering all essential elements. The participants were 47 real estate agents from different regions in Denmark and with varying experience, age, and background. 15 of the 47 real estate agents participated at the workshop in Vejen on June 6th, the other 32 at the workshop in Roskilde on June 12th. The participants were divided into two groups. Each group was by bus brought to the same 5 houses, one group at a time. In the bus, they received an information sheet for the next house; this sheet contained both general information on the municipality as well as key data on the houses’ characteristics and was supposed to convey important information on e.g. the location’s attractiveness to the – mostly non-local – real estate. The exact pieces on information that were provided for each house is listed in Table 4:
An example of such an information sheet (full version) is attached at the end of this chapter.

At each house, the real estate agents had approximately 20 minutes to assess the house price based on the given information and their own impression, observation and skills.

To ensure that the results were as unbiased and conclusive as possible, we applied the following experimental design:

The real estate agents were unaware that their task to assess house prices was, beyond being an element of the workshop, also part of an experiment. Consequently, there is no experimental bias – which means that the participants’ behaviour remains natural and unaffected by the experiment itself.

As a randomisation technique to divide the real estate agents into two groups at each workshop location, we used so called cluster randomisation, because it is favourable when working with small samples. Simple randomisation would mean that the allocation of participants into groups is fully left to chance. Although these groups will have identical characteristics on average, they will actually differ along some dimensions for small sample sizes. Exactly this problem advocates cluster randomisation. Cluster randomisation means that the allocation of participants into groups happens more deliberately: participants are divided into groups based on key variables. In our case, we made sure that all real estate agents with the same professional background (e.g. working in the same office) were divided equally into the two groups. The clear benefit of this method is that the resulting groups are more equal in its characteristics and therefore more easily comparable.

To ensure that the participants were motivated, a price was promised to the winner, who we defined as the one being (in total) closest to the assessment of the local real estate agent.
Additionally, it was announced that the assessments were non-anonymous and would be presented with names for the purpose of discussion in the workshops.

For each house, the real estate agents had circa 20 minutes to assess the value. As several participants stated, that was sufficient and in fact similar to the amount spent in actual assessments.

We avoided group dynamics by asking the participants to decide upon their assessment alone and unaided. We especially pointed out that discussing the house price with fellow participants or using the internet or similar aids is not permitted.

When choosing the houses, we ensured to cover a broad variety regarding size, value, and location. Variation in the energy labels was favourable, but only partly possible, since the real estate market is dominated by C- and D-labelled houses at the moment. Moreover, it was of importance that the energy label could be manipulated without causing scepticism among the participants.

As a robustness test, made sure to include jumps of two energy labels between the groups. If a 1-level jump in the energy label has an effect on the price, but this effect is so small that it is hard to detect, then it will become more evident for a jump of two labels. This robustness test makes sure that we do not overlook a small, but existing effect. The condition and characteristics of two houses in Roskilde allowed us to apply a jump of two energy labels between the groups. Larger jumps were not possible without endangering the believability in the information’s correctness.

We manipulated the energy label for one of the groups upwards and downwards for half the houses respectively to account for inherent, unobservable differences between the groups. In case that, for example, the real estate agents in group 2 were inherently more optimistic about sales prices and therefore gave constantly higher sales price estimates than group 1, then this bias would level out with our design. The resulting energy label effect remains unaffected.

An overview of the ten houses, their location, energy label levels and jumps as well as their value is given in the following table:
Table 5 Energy labels and house values

<table>
<thead>
<tr>
<th>House</th>
<th>Correct label</th>
<th>Label group 1</th>
<th>Label group 2</th>
<th>Jump in labels</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>V1</td>
<td>A2015</td>
<td>A2015</td>
<td>A2010</td>
<td>1</td>
<td>1,875,000</td>
</tr>
<tr>
<td>V2</td>
<td>C</td>
<td>C</td>
<td>B</td>
<td>1</td>
<td>1,495,000</td>
</tr>
<tr>
<td>V3</td>
<td>C</td>
<td>C</td>
<td>B</td>
<td>1</td>
<td>2,760,000</td>
</tr>
<tr>
<td>V4</td>
<td>E</td>
<td>E</td>
<td>D</td>
<td>1</td>
<td>975,000</td>
</tr>
<tr>
<td>V5</td>
<td>D</td>
<td>D</td>
<td>E</td>
<td>1</td>
<td>1,795,000</td>
</tr>
<tr>
<td>R1</td>
<td>C</td>
<td>C</td>
<td>D</td>
<td>1</td>
<td>2,400,000</td>
</tr>
<tr>
<td>R2</td>
<td>D</td>
<td>C</td>
<td>E</td>
<td>2</td>
<td>3,600,000</td>
</tr>
<tr>
<td>R3</td>
<td>D</td>
<td>D</td>
<td>E</td>
<td>1</td>
<td>2,795,000</td>
</tr>
<tr>
<td>R4</td>
<td>D</td>
<td>D</td>
<td>B</td>
<td>2</td>
<td>1,875,000</td>
</tr>
<tr>
<td>R5</td>
<td>D</td>
<td>D</td>
<td>E</td>
<td>1</td>
<td>4,650,000</td>
</tr>
</tbody>
</table>

Note: The houses V1-V5 were located in and around Vejen, R1-R5 in and around Roskilde.
The stated value is the house price assessed by the local responsible real estate agent.

Source: Copenhagen Economics.

3.3 Method / Testing for an effect

We apply a one-tailed, two-sample unequal variance t-test (Welch’s t-test) to test whether
the difference in energy label had a significant effect on the real estate agents’ assessment
of the house price.

Welch’s t-test, an adaption of the usual Student’s t-test, is designed to compare data from
two independent, non-identical groups or populations (“two-sample”) like our two groups
of real estate agents. We applied Welch’s t-test to all ten houses individually to determine
if the assessed house prices that occurred in each of the two groups are significantly differ-
ent from each other. Since the two groups have a small size\(^1\) and consist of different indi-
viduals, we must assume their variances to be non-identical (“unequal variance”).\(^2\)

The test statistic (t-value) in a Welch’s test is defined as:

\[
t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{N_1} + \frac{s_2^2}{N_2}}}
\]

with \(\bar{X}_1, \bar{X}_2\) being the sample means of sample 1 and 2,
\(s_1^2, s_2^2\) being the variances of sample 1 and 2, and
\(N_1, N_2\) being the sample size of sample 1 and 2.

---

\(^1\) Regarding a Welch’s t-test, a “small sample size” means less than 30 individuals.

\(^2\) In a second step, we test on a more aggregated scale for a difference between the average assessments of all houses with better energy labels compared to all those with poorer energy labels. For that comparison, we use a one-tailed, two-sample equal variance (instead of unequal variance) t-test. The shift towards equal variance results from the fact that we manipulated the energy label in both directions. The assessments of the houses with better and poorer labels respectively are therefore from a mix of group 1 and group 2 participants, and their variance can be assumed equal.
In other words, the t-value shows the difference between sample 1 and 2 expressed in units of the standard error, so measured relative to the variation in the sample. The higher the variance in the samples (and the smaller the sample size), the larger the standard error in the denominator and the smaller the t-value. The underlying sampling distribution of the test statistic is the t-distribution, which shows a reciprocal relationship between t- and p-value: a small t-value goes along with a large p-value. The p-value illustrates the probability of the occurrence the observed difference under the nil-hypothesis of both samples being equal; a small t-value and large p-value therefore cause the difference being insignificant and vice versa.

Next to the t-value, we need the degrees of freedom to look up the cumulative probability in the t-table and therewith to determine the p-value that indicates the significance level of the observed difference. The Welch test uses the Satterthaite-Welch adjustment to define the degrees of freedom:

\[
df = \frac{(\frac{s_1^2}{N_1} + \frac{s_2^2}{N_2})^2}{\frac{(s_1^2)}{N_1 - 1} + \frac{(s_2^2)}{N_2 - 1}}
\]

using the same notation as above.

The effect of a better energy label can be expected to be either insignificant or significantly positive, but not significantly negative. Consequently, we test for an effect in one direction of interest, which is why we apply a one-tailed instead of a two-tailed t-test.

### 3.4 Results

Does a better energy label give rise to a higher house price? Looking at the raw data, the physical experiment does not provide a clear answer to that question, neither on an aggregated nor on a single-house-level. We observe higher house price assessments for those houses with the better energy labels, but those differences are statistically insignificant.

The main cause for the insignificance and therefore the inconclusiveness of our data is the high variance of the assessments. To not overlook a positive energy label effect which is existent in, but not apparent from the data, we applied to methods to cleanse our data: firstly, we removed the outliers, secondly, we removed the houses which exhibited the highest variance.

Also after cleaning, the energy label effect remains positive, but statistically insignificant, cf. Figure 13. That the experiment does not prove an energy label effect on house prices, but it also does not disprove such an effect. The data is inconclusive, which does not allow a conclusion in neither direction.
This chapter is divided into three sections. In the first two sections, we show the results on an aggregated and on a more detailed, single-house level respectively. In the third section, we elaborate on the challenges posed by the data and how we attempt to overcome them.

The results on an aggregated level
Comparing the assessments of the ten houses, we see that the houses with the better labels were on average assigned a 157,000 DKK. higher price than those versions with the poorer label, cf. Figure 14. The direction of the effect – better label is assigned a higher price – is as expected, however, the difference is statistically insignificant.
Figure 14 Comparison of average house prices

Note: The left bar shows the grand mean of the "better-label-version" of the ten houses (mean of the ten houses’ means); the right bar shows the grand mean of the "poorer-label-version" respectively.

Source: Copenhagen Economics

The above figure shows the mean effect based on the mean assessments of the ten houses. The figure’s message, namely that there is a positive, but insignificant and therefore non-conclusive effect, is robust to other measurements of the average. Looking at the mean or median effect based on the mean or median assessments does not change the insignificance of the difference.

The results on a single-house-level
Looking at the single houses instead of the aggregated level neither supports any conclusion of an energy label effect; the data is inconclusive.

An intuitive illustration of the results is provided in Figure 15. The green arrows stand for a difference towards the expected direction; the shaded arrows with the question mark indicate that the difference is insignificant and that we therefore cannot be sure of the existence of that effect. For those houses with a fully filled arrow, the difference between the groups’ assessments is significant; the stars indicate the significance level.
The above illustration shows two problems that hinder conclusiveness:

*We observe “wrong-direction differences” for two out of the ten houses.* Only for six out of ten houses, we observe a higher average house price estimate in the group that was provided the better energy label. For two houses, the average estimate is indistinguishable in both groups, and for the remaining two houses, the real estate agents in the group with the better label estimated the house price lower, meaning we observe a “wrong-direction difference”. That means that only the assessment of six houses points to the conclusion of a positive energy label effect; a main reason for this and the following problem is the small sample size.

We observe insignificant differences for most houses. The second problem is that in most cases, the differences between the estimates of group 1 and 2 are statistically insignificant. Out of the assessments of the six houses which point to the conclusion of a positive energy label effect, only two show a significant difference. For the remaining four houses with a positive difference, the difference is statistically insignificant.

**High variance in the data – a problem to overcome**

The main reason for the inconclusive results is the high variance in the assessments. That means that the real estate agents made strongly diverging house price assessments both across and within groups. High sample variances automatically increase the p-value, especially in combination with small sample sizes, and therewith statistical insignificance. Noisy data like ours therefore complicates the detection of small effects.

The typical measure to describe the variation within a dataset is the standard deviation, which is the square root of the variance of a sample. A standard deviation close to zero means that all data points are close to the mean of the sample, whereas a large standard deviation indicates dispersed observations. For our ten houses’ price assessments, we observe large standard deviations; in few cases, the highest assessment is almost twice as high as the lowest one within the same group.

*We call the average assessments of the two groups “indistinguishable” when the two measures of average (median and mean) are contradicting, meaning when group 1 has the higher median but group 2 the higher mean or vice versa.*
The dispersion of the assessments, in particular how much the maximum and minimum assessment differ from the median, is illustrated in Figure 16:

**Figure 16 The variance of the house price assessments**

Note: The first graph shows the houses in Vejen (V1-V5), the second one the houses in Roskilde (R1-R5). For each house, there are two bars to account for the two groups. The pink markers illustrate the maximum and minimum assessment that was made.

Source: Copenhagen Economics

Reasons for the occurrence of such high variances in the assessments could possibly be the limiting factors like the fact that the participants were mainly non-locals.

Having been aware of the potential problem of group dynamics, we announced and later monitored that the real estate agents made their assessments individually and without talk-
ing to their colleagues; however, participants might have been influenced by their colleagues’ body language, countenance or by the questions they asked the owner, who was there at some houses.

The most crucial cause of high variance in our opinion is the fact that our participants were mainly non-locals. We expected our participants to be able to precisely assess the house price once they were given the key information about the area, but we seem to have underestimated the difficulty to do so. We are convinced we would face a significantly lower variance in the experimental data and would therewith be able to answer our question more clearly if the participants would have been locals, which was unfortunately not possible.

Hidden in the high variance, however, there might be a small, but existing energy label effect. To not overlook it, we apply two methods to reduce the noise in our data:

Firstly, we remove the outliers in each group and only look at those assessments around the median. We do that because we observe a few assessments for each house that are far away from the other estimates. This could be ascribed to inexperience, or could reflect real estate agents who might not have taken the task seriously. Their extremely high or low assessments then bias the average of the group and therefore the overall outcome.

Secondly, we look at the houses with the lowest variance only. Therewith we aim at remove all those houses which were, due to their characteristics, extraordinarily hard to assess, so that even experienced and committed real estate agents struggled to make a qualified assessment. The houses with the lowest variance are the where the assessment was less challenging, so the real estate agents were more agreed more on the price assessments. A potential energy label effect will be more visible for those houses.

a) Results after removing the outliers
Removing the outliers does not provide a clearer picture than the raw data. The hypothesis of the existence of an energy label effect can still not be corroborated. An overview on how removing the outliers changed the differences between the two groups is given in Figure 17:
We defined “outliers” in a relative rather than absolute way. That means we did not remove the assessments that were above or below particular threshold values, but that we removed a constant share of assessments. For each house, we neglected the one third of assessments, namely those that were furthest away from the median. That means we check if there is a significant and consistent energy label effect for the core two thirds of assessments per house.

The illustration above shows that the existence of outliers cannot have been the sole problem to blur a potential energy label effect. Removing the outliers reveals the before hidden positive effect for only one (namely the fourth house in Roskilde, R4) out of the four houses that had yellow or red arrow before. Further improvements can be seen for house V2 and R2, which show a (more) significant positive difference between the two groups when only looking at the core two thirds of the assessments. For the houses V5 and R3, however, we see a negative development; that means that in that cases, the outliers actually helped hiding a negative difference which contradicts the energy label effect.

b) Results for the houses with the lowest variance
Looking at the five houses with a low variance might hint towards, but does neither prove the existence of an energy label effect.

The selection of houses is dominated by green arrows (meaning that the difference has the right direction), but only one shows a significant, positive effect for both the raw and the cleansed data, namely R2, cf. Figure 18. Among the houses that were removed where many observations which did not support an energy label effect (V1, R3, R4 for the raw data) –

---

4 An exception was made if there was a large gap between a highly compact core of assessments on the one hand and the outliers on the other hand; in that case, we deviated slightly from our $\frac{1}{3} - \frac{2}{3}$ rule and identified (at maximum) one more or one less assessments as outliers.
which is good news – but also two that actually showed a positive an in one case even posi-
tive and highly significant difference (V4 and R3). Moreover, the fifth house in Vejen is
part of the selected low-variance houses, but shows no (for the raw data) and a negative
(for the cleansed data) difference.

Figure 18: The energy label effect for low-variance-houses

Raw data:

![Diagram]

Cleansed data:

![Diagram]

Note: Significance levels: '. ' =10%, '* =5%, '***' =1%, '****' =0.1%.

Source: Copenhagen Economics

The average standard deviation expressed as a percentage of the median assessment was
the criterion we used to select the houses with the low variance. We used percentages of the
median instead of the absolute standard deviation in order to not give an advantage to those
houses with lower values.

The standard deviation for each group and house and their average as well as the respective
percentage values – the selection criterion which Figure 18 is based on – are shown in Table
6. The houses are sorted after the average percentage deviation; the upper five in pink are
the selected low-variance houses.
### Table 6 Variances in the house price assessments

<table>
<thead>
<tr>
<th>house</th>
<th>group 1</th>
<th>group 2</th>
<th>mean</th>
<th>group 1 as percentage</th>
<th>group 2 as percentage</th>
<th>mean as percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>R2</td>
<td>287,015</td>
<td>449,239</td>
<td>368,127</td>
<td>8%</td>
<td>13%</td>
<td>11%</td>
</tr>
<tr>
<td>V5</td>
<td>188,754</td>
<td>193,010</td>
<td>190,882</td>
<td>11%</td>
<td>11%</td>
<td>11%</td>
</tr>
<tr>
<td>V3</td>
<td>281,761</td>
<td>394,882</td>
<td>338,322</td>
<td>11%</td>
<td>14%</td>
<td>12%</td>
</tr>
<tr>
<td>R1</td>
<td>304,327</td>
<td>250,512</td>
<td>277,419</td>
<td>14%</td>
<td>12%</td>
<td>13%</td>
</tr>
<tr>
<td>V2</td>
<td>254,169</td>
<td>216,458</td>
<td>235,313</td>
<td>17%</td>
<td>13%</td>
<td>15%</td>
</tr>
<tr>
<td>R4</td>
<td>353,258</td>
<td>250,072</td>
<td>301,665</td>
<td>19%</td>
<td>13%</td>
<td>16%</td>
</tr>
<tr>
<td>V4</td>
<td>106,250</td>
<td>178,462</td>
<td>142,356</td>
<td>13%</td>
<td>21%</td>
<td>17%</td>
</tr>
<tr>
<td>V1</td>
<td>361,385</td>
<td>223,899</td>
<td>292,642</td>
<td>22%</td>
<td>11%</td>
<td>17%</td>
</tr>
<tr>
<td>R3</td>
<td>376,407</td>
<td>359,962</td>
<td>368,185</td>
<td>19%</td>
<td>18%</td>
<td>19%</td>
</tr>
<tr>
<td>R5</td>
<td>942,532</td>
<td>763,971</td>
<td>853,251</td>
<td>19%</td>
<td>21%</td>
<td>20%</td>
</tr>
</tbody>
</table>

**Note:** The houses are sorted after the last column, the mean of the standard deviation expressed as a percentage.

**Source:** Copenhagen Economics
Chapter 4

The conjoint analysis

4.1 Why a conjoint analysis?

The price that consumers are willing to pay for a product reflects how they value this product – or more precisely, it reflects how they value the particular bundle of attributes offered by this product. In the context of the real estate market, this ‘product’ is a house, and people will – wittingly or unwittingly – assess what a particular house’s size, architecture, garden, location, condition, energy label, and so forth is worth to them.

Since the market price of a particular product always reflects the willingness to pay (WTP) for the whole bundle of its attributes, the market price does not reveal the ascribed value of one particular attribute within that bundle on that price. In our case, we are interested in the effect of the energy label on the house price, but looking at the prices of houses with different energy labels is not sufficiently informative. A high WTP for a house could reflect the consumers’ appreciation of its high energy label, but also its architecture, location, garden, or any other attribute of the bundle. To understand the role of energy labels, we need to “disassemble” the customers’ valuation of a house into its single elements, the valuation of the different attributes.

Choice modelling methods allow us to do that. By asking respondents to rank or choose between different houses, we gain insight into which element of the bundle is important to them. Choice modelling methods reveal the partial preferences of and relations between attributes. They measure the relative values of different characteristics and, by that, provide detailed understanding on how people choose among various competing alternatives.

There are four different types of choice modelling methods. Each type has different benefits and drawbacks and therefore serves a particular purpose. In a conjoint analysis (also called choice experiment), respondents choose their most preferred alternative from a choice set with two or more options. The same is done for paired comparisons, but here, participants additionally indicate how strongly they prefer the chosen alternative over the other(s) on a predefined numeric or semantic scale. In a contingent ranking, respondents are presented three or more alternatives which they rank from the most to least preferred. Lastly, contingent rating comprises the respondents to rate different alternatives on a predefined semantic or numeric scale. Alternatives are not compared in this method, only one is presented at a time.

An overview on the four different choice modelling methods is provided in Figure 19.

---

5 Due to different tastes, incomes and backgrounds, people’s preferences for the same attribute differ. For a young family, a location close to a school might be a good thing, while it might be much less desirable for an elderly couple. People with a car might value a garage next to the house, while those without a car may put more weight on living close to a bus or train station. In our analysis, we do not focus on this individual level, but examine preferences on an aggregated market level instead. That provides us with the insight into how the overall population value these attributes.
For the type of analysis we will undertake, the conjoint analysis is the most appropriate among the choice modelling methods.

The conjoint analysis’ main advantage is its simplicity regarding the implementation: Choosing one out of two alternatives is a straightforward task and applied in our everyday decisions. All other choice modelling methods would involve ranking or rating of the houses, which requires (1) detailed instructions as well as (2) a longer phase of considerations for the participants compared to simple choices. Consequently, a choice modelling method that includes rating or ranking usually forces the experimenter to reduce the number of choice sets to not overwhelm the participant. Compared to contingent rating, contingent ranking and paired comparisons, the conjoint analysis allows us therefore to expand our data collection concerning the number of choice sets. Especially with a limited number of participants like in our case, that is crucial to ensure significant results.

4.2 Idea in combination with the other experiments
The conjoint analysis is an integral part of our package of experiments. Its main goal being to reveal how people rank the energy label in relation to other house characteristics, the conjoint analysis approaches the examination of the effect of energy labels on house prices from a different angle than the net-based and physical experiment.

The conjoint analysis moreover contributes to a broad analysis by covering a more hypothetical and simplified approach than the other experiments. In the conjoint analysis, participants are confronted with sufficient, but limited information. All characteristics of the houses being equal except for a few key ones allows us to break down the value of the different characteristics, including the energy label.
4.3 Setting and design of the survey
This chapter is divided into two parts: the first sub-chapter illustrates the setting, meaning the practical execution of the experiment, while the second sub-chapter explains its design.

General setting
Our conjoint analysis survey that we gave out to 47 real estate agents consisted of one page of clarifications and instructions followed by 24 sets of two houses each. For each set, participants had to choose the more valuable house. The participants were the same real estate agents as for the physical experiment described in chapter 3. They all participated at a workshop conducted in Vejen and Roskilde respectively, organised by the Dansk Ejendomsmæglerforening (The Danish Association of Chartered Estate Agents). The focus of the workshop was the assessment of house values, so the conjoint analysis fit in well without being thematically surprising.

The atmosphere was impartial and motivated, since energy labels and their importance had not been addressed before. The seriousness of the conjoint analysis was underlined by the setting being an official workshop, as well as non-anonymity and a limited time of 15 minutes for finishing the experiment – which pilot studies that we conducted beforehand had proven to be sufficient. Based on these circumstances, the participants made deliberate choices and the experiment appears to be unbiased.

Design of the choice sets
The 24 so-called ‘choice sets’ of the survey were of the following general form:

| 190 kvm hus | vs. | 160 kvm hus |
| 700 kvm grund | | 900 kvm grund |
| moderat stand | | god stand |
| energimærkning B | | energimærkning E |
| ville blive solgt for den højeste pris | |

As it can be seen from the example above, the real estate agents faced two alternative houses, of which they were asked to choose the one which will be sold for the higher price, so the one of the higher market value. They are not asked for their personal preferences, and we point that out in particular in the instructions so ensure that our results refer to market values. This is important, since we aim at assessing the effect of energy labels on house prices for the real estate market, meaning for an aggregated level. Looking at the individual level instead would lead to results, too, but we would not be allowed to generalise them. By asking real estate agents for market values, we get representative, generalizable results for the real estate market, based on the knowledge of those who know this market best.

Each of the 24 choice sets consists of two houses, the so-called ‘alternatives’. Each alternative depicts a house, which is characterised by four so-called ‘attributes’ (egenskaber);
those attributes in turn have several levels (niveauer). The attributes and levels we used for our Conjoint Analysis are shown in Table 7:

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Levels</th>
</tr>
</thead>
<tbody>
<tr>
<td>House size</td>
<td>100 sqm, 130 sqm, 160 sqm, 190 sqm</td>
</tr>
<tr>
<td>Site size</td>
<td>600 sqm, 700 sqm, 800 sqm, 900 sqm</td>
</tr>
<tr>
<td>Condition of the house</td>
<td>moderate, good</td>
</tr>
<tr>
<td>Energy label</td>
<td>B, C, D, E, F, G</td>
</tr>
</tbody>
</table>

Source: Copenhagen Economics

An efficient design of the conjoint analysis is crucial for sound results. To find the best experimental design, a range of important decisions regarding the details of the experiment has to be made. The following sub-chapters provide the explanations and reasoning behind four crucial decisions and give an insight into the processes of deciding upon the attributes (a-b), levels (c), and the composition of the choice sets (d).

**a) Deciding upon the number of attributes – as many as possible, as few as necessary**

For our Conjoint Analysis, we worked with four attributes. To find that number, we had to find the balance in a trade-off: On the one hand, we want to mimic a real-life choice to put the real estate agents into a familiar situation. Since each house is characterized by countless attributes, including more entails the choice to be less artificial. On the other hand, there are two arguments that inhibit us from including many characteristics:

*Firstly*, people are easily overwhelmed with a choice if there are many facts to base the decision upon. Their choices are then likely to be irrational or random. The recommendable number of attributes differs from case to case, but in general, there should not be much more than a handful of attributes to ensure that the participants are able to take all given information into account when making their choices.

*Secondly*, an increasing number of attributes generally goes along with decreasingly robust results. That is because there is a maximum amount of choice sets we can ask people to answer, but with more attributes, the amount of choice sets we need to ask to gain trustworthy results increases.

We are limited regarding the amount of choice sets we can give to the participant, because after too many choices, people’s concentration and motivation decreases and they tend to get tired, bored or annoyed. Their choices will then no longer be deliberate, but irrational and random, which makes the data useless for an analysis. At the same time, the number of all possible combinations of attributes and levels, the so-called full factorial design, increases exponentially in attributes:
In our case, we decided to use four attributes with 2, 4, 4 and 6 levels respectively. That means our full factorial design sums up to:

\[ 2^1 \times 4^2 \times 6^1 = 192 \text{ alternatives} \]

Those 192 alternatives correspond to 96 choices between two alternatives each. Based on a pilot study we conducted, we found that 24 choice sets – so 48 alternatives – is the maximum amount that the real estate agents could answer intently. These 48 alternatives out of the possible 92 are called a fractional factorial design. Since we limited our attributes and levels to a reasonable amount, this fractional factorial design is a sufficiently large share of the full factorial design to allow for sound and representative results. This case is depicted in the illustration to the right in Figure 20 below. In the case of choosing too many attributes, the full factorial design will surge: because of the exponential relationship, already two more attributes with four levels each will suffice to lead to a rise the possible combinations from 192 to 3,072. The fraction that needs to be applied to provide sound results will surge accordingly. The actual fractional factorial design of the survey will then not be sufficiently large, as depicted in the left illustration in Figure 20.

**Figure 20 The perfect size of the fractional factorial design**

The best size of the fractional factorial design is not a particular number that can be calculated; neither is there an exact number of choice sets that must not be exceeded. An appropriate size is rather found by gauging the explained pro’s and con’s of increasing the fractional factorial design and is mainly based on one’s experience; the D-efficiency indicator which is calculated by R provides additional information. The perfect size of the fractional factorial design depends on the full factorial design and the complexity of the task; it is independent of the number of participants (and therefore the number of observations).
b) **Deciding upon the concrete selection of attributes**

Being aware of the importance of including just a few attributes raises the question: *Which ones to pick?*

Without any doubt, the energy label must be one of the attributes, since it is their effect that we examine. Next to the energy label, we decided to use the size of the house, the size of the site and a quality indicator, which we called the condition of the house.

Considerate thinking combined with findings from the academic world helped us finding the attributes that are best for our experimental design. The following four requirements should be fulfilled:

1. **The chosen attributes must be relevant** for making the choice. The information provided must allow the participant to make a deliberate decision. That is essential to mimic a real-life situation as well as possible.

   In our case of the real estate market, we must *provide crucial information* like the size of the house or the location. If we asked the real estate agents to choose the more valuable house among two of which they lack this information – it is likely that they either would refuse to do so, or make a random choice, since they do not have the necessary facts to base their decision on.

2. **None of the chosen attributes must dominate** the choice.

   Dominating the choice means that one characteristic is far more important than the others, so that the choice is solely based on this one attribute – all the others will then show now significant effect, and the purpose of the analysis is lost. In our case of the real estate market, we consider the location of the house such a dominating attribute. The differences in popularity of the municipalities are too explicit. It is likely that a real estate agent will always choose a house in Copenhagen or Gentofte over one in Lolland or Æro, where the square meter prices reach only about one sixth of the former (boliga.dk).

   Since leaving the location of the houses out would contradict the argument in the first bullet point, we include it, but as a given fact for all houses, so that the choice is independent of it.

3. **The attribute levels between the different attributes must be uncorrelated.** This requirement is called *orthogonality* and is one of Huber and Zwerina’s (1996) criteria for an efficient experimental design, to which we will come back in the next section.

   The requirement of orthogonality would for example be violated if we included both the energy label and the condition of the roof. A roof in a bad condition decreases the energy efficiency of the house, and vice versa, which means that the energy label is no longer independent of the other attributes. Correlated attributes like these generate the following dilemma:
Either one ignores this relationship and places the affected attributes’ levels fully independently. That would mean that the alternatives would include houses with a roof in a bad condition but energy label A or, the other way round, houses with a roof in a good condition but energy label G. This method acts against common sense, and consequently, the participants will be confused and make nonsensical and undeliberate choices.

Alternatively, one can consider the existing relationship between the two attributes and place the attributes’ levels in a way that does not violate common sense. That would mean that an energy label A automatically excludes the possibility of having a roof in a bad condition. Applying this method allows participants to make deliberate choices, but entails an econometric problem: effects can then not be tracked back to their actual trigger. Participants might have chosen a particular house because of the good label or because of the good condition of the roof – the effects of the two correlated attributes cannot be separated in the analysis.

4. Attributes that leave little to no room for interpretation are favourable. As a rule of thumb, one can say that attributes with numeric levels are preferred over attributes with semantic levels.

   Numeric levels means that the attribute takes on numeric values, like 100, 130, 160 and 190 square meters for the size of the house in our case. Those values have the advantage that all participants will understand them in exactly the same way; there is no room for interpretation. Semantic levels means that the attribute takes on descriptive words, for example “small”, “rather small”, “rather large” and “large” instead of the square meter values. If not guided by a detailed explanation what exactly those words mean, participants will interpret them differently. The different interpretation among participants will result in an effect that the model cannot explain. The analysis will be biased.

   In case semantic levels are necessary, it is crucial to include a detailed explanation of what those descriptive words mean exactly – that removes the room for interpretation and ensures that all participants understand the attributes and its levels in the same way.

With our selection of attributes, we meet all the requirements above.

The size of house and site as well as an indication of the condition of the house are definitely among those characteristics that determine the value of the property.

So does the location of the house – but as mentioned above, this characteristic would be dominating the others. It is therefore important to include information on the location, but without letting it influence the choice. We strike this balance by placing all 48 houses in our 24 choice sets in the same location, namely in Bagsværd in the periphery of Copenhagen. In the instructions of our survey, we explicitly point out that the two houses of each choice
set should be considered as neighbouring houses. By making the location a neutral characteristic in our choices, we avert any unwanted, distortive effect that the location might have caused when not treated properly.

Meeting the requirement of orthogonality significantly influenced the selection of attributes. Since we examine the effect of the energy labels, it was of the utmost importance to include no other attribute that is correlated with the labels. That means we were not allowed to include the build year, a characteristic that ranges among the important ones when it comes to assessing the value of a house – nor could we apply the method of including it as an invariable fact, as we did for the location. The reason is that including a fixed build year would mean that this year applied to all energy labels – but it is rather unlikely that very new houses have a label worse than D, just as it is unlikely that very old houses have a label better than C. There is no build year that would apply to all labels without causing confusion among the participants. We therefore decided to point out that each two houses of one choice set were built in the same year. By mentioning that explicitly, we are confident that our participants do not let their own, differing assumptions on the build year influence their choice. Fulfilling orthogonality also hindered us from including information on the windows, roof or wall insulation, elements that are highly correlated with the energy label. Since those facts must rather be considered elements of the energy label instead of correlated attributes, this is a minor issue: those elements are automatically included by including the energy label.

To make sure that the real estate agents do not perceive the condition of the house and the energy label as correlated, we defined the condition as related to the kitchen and bathroom. Additionally, we underlined in the instructions that these two attributes are uncorrelated. Finally, we also meet the requirement of little room for interpretation. Three of our four attributes have numeric levels; the one with semantic levels is the condition of the house, which can be either ‘good’ or ‘moderate’ in our survey. To make sure that every one of the 47 real estate agents has the same understanding of what ‘good’ and ‘moderate’ means, we included a comprehensive and detailed explanation in the instructions. In this explanation, we use tangible examples about the condition of kitchen and bathroom, and state that the condition of all other, non-mentioned elements are identical.

Following these four requirements, we moreover decided not to use any pictures in our survey. There is a prevalent risk of the pictures being distortive: the appearance of houses is likely to be perceived correlated with the energy label, and additionally, pictures leave room for interpretation, since they trigger different ideas and assumptions about the elements that are not shown.

c) Deciding upon the levels and the experimental design

After having decided upon which attributes to use, we needed to find appropriate levels for them. This task requires common sense and an insight into the market, as well as the consideration of some criteria: For the same reasons as for the attributes, the levels should be reasonable, not too many in number and not causing a dominance.

For our attributes, we chose a 4-4-2-6 design with the following levels (cf. Table 8):
Table 8 Details on the explanatory variables

<table>
<thead>
<tr>
<th>Attribute</th>
<th>No. of levels</th>
<th>Levels</th>
<th>Variable type</th>
</tr>
</thead>
<tbody>
<tr>
<td>House size</td>
<td>4</td>
<td>100 sqm, 130 sqm, 160 sqm, 190 sqm</td>
<td>Discrete</td>
</tr>
<tr>
<td>Site size</td>
<td>4</td>
<td>600 sqm, 700 sqm, 800 sqm, 900 sqm</td>
<td>Discrete</td>
</tr>
<tr>
<td>Condition of the house</td>
<td>2</td>
<td>moderate, good</td>
<td>Dummy</td>
</tr>
<tr>
<td>Energy label</td>
<td>6</td>
<td>B, C, D, E, F, G</td>
<td>Dummy</td>
</tr>
</tbody>
</table>

Source: Copenhagen Economics

The house size as well as the site size are discrete variables; that means we will be able to calculate the effect per square meter. We chose to use four levels, because fewer levels would only allow for a dummy interpretation.

For the condition of the house, we chose to use the levels ‘moderate’ and ‘good’. We decided not to use ‘bad’ or ‘poor’, because it might cause the condition to be a dominating attribute. People might not choose a house in a ‘bad condition’ over another one, regardless of the other attributes.

Regarding the labels, we decided to use the energy labels B to G as dummy variables. We left out energy label A to not cause the full factorial design to grow too much, but also to ease level balance, a criterion for an efficient experimental design, that means that the levels of all attributes should have a low common denominator. We will come back to this criterion in the next section on the experimental design.

d) The experimental design of the choice sets

To produce results that are reliable, the experimental design of our survey needs to be efficient. To achieve efficiency, we ensured that the design of our survey meets the following five criteria (partly based on Huber and Zwerina 1996):

1. **Orthogonality**
   The attribute levels between the different attributes must be uncorrelated.
   As explained in the section b), our design meets orthogonality.

2. **Level balance**
   The levels of an attribute should appear with equal frequencies in the survey to ensure that the results are of a constant robustness.
   In our survey 48 alternatives, each level of house and site size appears in exactly 12 alternatives (1/4 of 48), moderate and good condition appear in 24 alternatives each (1/2 of 48), and each energy label in exactly 8 alternatives (1/6 of 48). We perfectly meet level balance.

3. **Utility balance**
   Within one choice set, the two alternatives are close in utility, so that the participants trade at the margin.
We do not have any dominating attributes in our survey. Moreover, we paired the alternatives in a way so that there is always a trade-off, meaning one alternative can at a maximum be better than the other alternative in three attributes – never in all four.

4. Minimal overlap
Within one choice set, the levels of attributes differ across alternatives. This criterion is automatically met in a 4-4-2-6 design as ours, for which same or very similar alternatives in one choice set are highly unlikely.

5. No ordering bias
The experimental design must ensure that the order in which the attributes are presented does not distort the results. People tend to unwittingly ascribe more weight to those attributes that are named first. To avoid this so-called ordering bias without confronting the participants with a confusing mix of differently ordered choice sets, we controlled for that bias across participants. Our survey consisted of four versions, each of them with a different ordering of the attributes, which we gave out to one fourth of the participants respectively. Thereby, all attributes appear at each position with the same frequency.

An example of the different versions of one alternative is given in the following Figure 21.

Figure 21 Differently ordered versions of the same alternative

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>190 kmv hus</td>
<td>energimærkning B</td>
<td>700 kmv grund</td>
<td>moderat stand</td>
</tr>
<tr>
<td>700 kmv grund</td>
<td>700 kmv grund</td>
<td>190 kmv hus</td>
<td>energimærkning B</td>
</tr>
<tr>
<td>moderat stand</td>
<td>energimærkning B</td>
<td>700 kmv grund</td>
<td></td>
</tr>
<tr>
<td>energimærkning B</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: Copenhagen Economics

4.4 The model
For the analysis of our data, we use a *conditional logit model*. The conditional logit is the typical model to use for all those Choice Modelling methods that focus on choice attributes and not characteristics of the individual that makes the choice.

In simple words, the model compares the characteristics of all the preferred houses to those of all the not chosen houses and from that calculates how important which characteristic has been to the participants. As typical for logit models, the model works with probabilities to describe such influences of characteristics on the choice of participants.
Our dependent variable is the choice variable, which has a binary output of either ‘1’ or ‘0’. This choice variable exists for each alternative, so for each house; ‘1’ means that it was chosen to be the more valuable house, ‘0’ means it was not chosen.

As all logit models, the conditional logit model is based on the cumulative distribution function (CDF) of a logit function. The CDF describes the probability that our dependent variable (the choice variable) will be found to have value less than or equal to \( x \) and is shown in Figure 22.

**Figure 22 The cumulative distribution function in a logit model**

![Cumulative distribution function](image)

Source: Copenhagen Economics

For each house that was chosen (not chosen) in our survey, the model notes the dependent variable as positioned right (left) of \( x \).

For each participant and each alternative, the model connects the position of the dependent variable with the levels of the attributes that led to that choice. Therewith, it calculates to which extend which of the explanatory variables influenced the odds (probability) of making that choice. This effect of each attribute on the odds of choosing a house is what we find in our results table.

One important underlying assumption for the model is utility maximising behaviour. Each alternative provides – due to its particular attributes and levels – a certain utility to the participant, who then chooses the one from which he derives the higher utility. Generalised for all choice sets, each participant faces the choice between the two alternatives A and B and therefore between the two utilities:
with $U$ being the total utility derived from that alternative (A or B), house, site, condition and label being our explanatory variables (with differing levels across alternatives) and the betas being an indication for their importance with regards to the derived utility. Those betas are, however, not shown in the regression output of a logit model, which works with more complex, probability-based calculations. $k$ stands for an alternative specific constant, which captures the average effect on utility of all factors that are not included in the model, and $\varepsilon$ is the error term.

4.5 Results
The results of a logit regression are rather complicated to interpret, which is why we start with the most intuitive and common interpretation – namely in terms of trade-offs – before we go into details explaining the regression results.

When looking at trade-offs that people make between different attributes, we see two characteristics for which the participants’ utility remains constant – i.e. they are indifferent of the two options. We can for example look at how many square meters of the house people would trade off for a condition that is ‘good’ instead of ‘moderate’, or for an energy label B instead of C. Technically, we can compute those trade-offs for all combinations of our attributes. However, it is most interesting to look at the trade-offs between energy labels and house size, because (1) compared to the site size, the house size has a highly significant estimate and (2) compared to the condition, the house size is a non-dummy variable, so trade-offs can be interpreted suitably.

The trade-offs are calculated as follows:

$$
\text{trade-off}_{\text{house size, energy label}} = \frac{\text{estimate energy label}}{\text{estimate house size}}
$$

The trade-offs between the energy labels and the house size are shown in Figure 23:
The relative effects are equal to the marginal changes in odds, and so are the significance levels. The benefit of this illustration is that we can interpret the energy label effect in a much more tangible way. The pink areas illustrate the trade-offs from jumping one energy label upwards. Generalizing our findings, people would for example be willing to trade off 15 square meters of a house for an upgrade from label F to E, or 11.3 square meters for an upgrade from C to B. The total height of the columns and the numbers above them show the trade-off for the total difference from G to each of the labels respectively. For an upgrade from energy label G to C for example, people are willing to trade off 35.3 square meters of the house.

Back to the original econometric results: running the logit model for our collected data with the choice as a dependent variable gives the following results (cf. Table 9):

**Table 9 Logit regression results**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>Significance</th>
<th>Marginal change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-0.355</td>
<td>0.102</td>
<td>***</td>
<td></td>
</tr>
<tr>
<td>House size</td>
<td>0.044</td>
<td>0.004</td>
<td>***</td>
<td>0.05</td>
</tr>
<tr>
<td>Site size</td>
<td>0.001</td>
<td>0.001</td>
<td>.</td>
<td>0.00</td>
</tr>
<tr>
<td>Condition</td>
<td>0.984</td>
<td>0.168</td>
<td>***</td>
<td>1.67</td>
</tr>
<tr>
<td>Label F</td>
<td>0.108</td>
<td>0.207</td>
<td></td>
<td>0.11</td>
</tr>
<tr>
<td>Label E</td>
<td>0.770</td>
<td>0.246</td>
<td>**</td>
<td>1.16</td>
</tr>
<tr>
<td>Label D</td>
<td>1.321</td>
<td>0.261</td>
<td>***</td>
<td>2.75</td>
</tr>
<tr>
<td>Label C</td>
<td>1.554</td>
<td>0.319</td>
<td>***</td>
<td>3.73</td>
</tr>
<tr>
<td>Label B</td>
<td>2.053</td>
<td>0.325</td>
<td>***</td>
<td>6.79</td>
</tr>
</tbody>
</table>

Note: Significance levels: ‘.’ = 10%, ‘*’ = 5%, ‘**’ = 1%, ‘***’ = 0.1%.

Source: Copenhagen Economics
The estimates itself of a logit model cannot be interpreted directly – but we can look at the signs and compare level estimates within the same attribute.

A positive number indicates a positive effect on choosing the house. We can see that it becomes more likely to choose a house with each additional square meter of house and site, a better condition (base dummy: ‘moderate’) and a better label (base dummy: ‘G’).

The intercept in a choice model has no interpretative value, but is typically small or negative if there are dummy variables involved. Both the condition and the energy label are dummy variables in our case. The bases of those dummies affect the intercept, and since our base dummies ‘moderate condition’ and ‘energy label G’ depict rather unwanted characteristics of a house, we see a negative relationship.

Looking at the column showing the significance of our results we can see that our experimental design produced very robust results. Most of the variables are highly significant. The only insignificant variable is ‘label F’, which still has a positive estimate. That suggests that the model found a positive effect on the choice from jumping from label G to label F, but the indications are not strong enough to be sure of the existence of this effect.

The most informative column is the one to the very right, which shows the marginal change in odds. As typically for logit models, they are calculated as followed:

\[
\text{Marginal change in odds} = e^{\text{estimate}} - 1
\]

Those marginal changes can be interpreted as percentages. Each additional square meter of a house for example increases the odds of it being chosen by 5%. The effect of an additional square meter of the size is still significant on a 10%-level, but the change in odds is smaller than 0.5%, so the effect is in fact negligible. A jump from moderate to good condition leads to an increase in the odds of it being chosen by 167%.

The change in odds caused by jumps in energy labels are particularly interesting. Since the base dummy for the energy label is ‘G’, the marginal change in odds is interpreted as the effect from jumping from label G to the respective higher label. For a jump of only one label, from G to F, the model suggests an increase in odds of choosing the house of 11%, but since the result is insignificant, we cannot be sure of that effect. However, as soon as we look at a jump of more than one level, the effect becomes both highly significant and remarkably large. Jumping from G to E for increases the odds of choosing the house by 116%, and that change in odds grows up to 679% for a jump from G to B.

Those percentage values seem very high, but are reasonable and must not be mistaken as an indication for certainty that a house will be chosen. We have to keep in mind that the base dummy is label G; the initial odds of choosing a house with label G are likely to be very low. These small initial odds increase by 679% when jumping up to label B – the result is a remarkably higher probability, but no certainty whatsoever.
Since we worked with dummy variables for the energy labels, we can use any of the six labels as base dummy and therefore are able to examine each single jump within the range of B to G individually.

The marginal changes in odds and their significance for all possible jumps within the range of label B to G are shown in Table 10. For each cell, the label indicated left in the same row is the initial label; the label on top in the same column is the level we jump to. The pink cells indicate all jumps of one label respectively.

### Table 10 Marginal changes in odds

<table>
<thead>
<tr>
<th>Jump from to</th>
<th>Label G</th>
<th>Label F</th>
<th>Label E</th>
<th>Label D</th>
<th>Label C</th>
<th>Label B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Label G</td>
<td></td>
<td>11%</td>
<td>116%</td>
<td>275%</td>
<td>373%</td>
<td>679%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.60</td>
<td>**</td>
<td>***</td>
<td>***</td>
<td>***</td>
</tr>
<tr>
<td>Label F</td>
<td>-10%</td>
<td>94%</td>
<td>236%</td>
<td>325%</td>
<td>599%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>**</td>
<td></td>
<td>***</td>
<td>***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Label E</td>
<td>-54%</td>
<td>-48%</td>
<td>73%</td>
<td>119%</td>
<td>261%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>**</td>
<td>**</td>
<td>***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Label D</td>
<td>-73%</td>
<td>-70%</td>
<td>-42%</td>
<td>26%</td>
<td>108%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>***</td>
<td>***</td>
<td>***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Label C</td>
<td>-79%</td>
<td>-76%</td>
<td>-54%</td>
<td>-21%</td>
<td>65%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>***</td>
<td>***</td>
<td>***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Label B</td>
<td>-87%</td>
<td>-86%</td>
<td>-72%</td>
<td>-52%</td>
<td>-39%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>***</td>
<td>***</td>
<td>***</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Significance levels: ‘.’ = 10%, ‘*’ = 5%, ‘**’ = 1%, ‘***’ = 0.1%. The two numbers 0.60 and 0.25 stand for the p-values of those jumps — that means these jumps are insignificant. The p-value is supposed to provide an idea about ‘how insignificant’.

Source: Copenhagen Economics

The above overview provides a range of insights. The following list is a selection of those:

- Each upward jump increases the odds of choosing the house.
- The positive effect is not constant across the labels; some jumps have a larger effect than others do.
- The increasing effect is consistent across the number of jumps. At any starting level, a jump of two labels always leads to a larger effect than a jump of one label.
- Three out of the five 1-label-jumps and all jumps of more than labels have a significant effect.

The effects of all jumps of one label (those that were marked as pink in the table above) are illustrated in Figure 24 in a diagrammed style:
Figure 24 The effect of the energy label on the choice

The chart gives an overview on the relative effects and shows that the “energy label effect” depends on the initial level. According to the data from our participating real estate agents, a jump from G to F (11%) as well as from D to C (26%) does not make a big difference for the house value. The labels B, D and especially E on the other hand seem to function as a threshold; energy efficiency improvement that upgrades the houses' label to B, D or E increases the odds of choosing that house by 65 to 94%.

Note: The baseline is energy label G.
Source: Copenhagen Economics
Appendix A

House pictures in web-based experiment

2A

Beliggenhed: Køge
Boligareal: 127 m²
Grundareal: 445 m²
Rum: 4
Antal plan: 2
Byggeår: 1957
Energimærke: X


2B

Beliggenhed: Køge
Boligareal: 123 m²
Grundareal: 437 m²
Rum: 4
Antal plan: 2
Byggeår: 1954
Energimærke: X


3A

Beliggenhed: Kolding
Boligareal: 204 m²
Grundareal: 432 m²
Rum: 4
Antal plan: 2
Byggeår: 1925
Energimærke: X

**3B**

Beliggenhed: Kolding  
Boligareal: 198 m²  
Grundareal: 439 m²  
Rum: 4  
Antal plan: 2  
Byggeår: 1929  
Energimærke:  


**4A**

Beliggenhed: Sorgenfri  
Boligareal: 135 m²  
Grundareal: 448 m²  
Rum: 4  
Antal plan: 2  
Byggeår: 1926  
Energimærke: X  


**4B**

Beliggenhed: Sorgenfri  
Boligareal: 140 m²  
Grundareal: 452 m²  
Rum: 4  
Antal plan: 2  
Byggeår: 1932  
Energimærke: X  

Lækker åbent køkken med adgang til to meget flotte, lyse stuer. 1. salen indeholder to store, lækre værelser samt et stort pænt badeværelse holdt i lyse, flotte klinker. Sidste renovering blev foretaget i 2009.


Deloitte (2012) Analyse af effekten af energioptimering af ejerboliger

EDC (2012) Godt energimærke sælger din bolig


SBi (2013a): Sammenhæng mellem energimærkning og salgspris Statens Byggeforskningsinstitut in corporation with University of Aalborg, 2013(06)

SBi (2013b): Cost-optimal levels of minimum energy performance requirements in the Danish Building Regulations, Statens Byggeforskningsinstitut 2013(25)

Realkredit Danmark (2013) Energimærkning og boligøkonomien