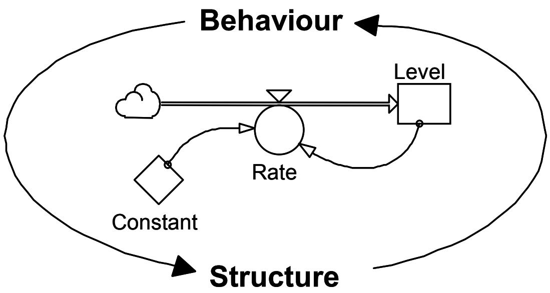
WORKING PAPERS IN SYSTEM DYNAMICS

WPSD 4/03 **ISSN 1503-4860**

Estimating customer utility of energy efficiency standards for refrigerators

by

Erling Moxnes



The System Dynamics Group

Department of Information Science
University of Bergen, Norway
Postbox 7800, N-5020 Bergen, Norway

Estimating customer utility of energy efficiency standards for refrigerators

By Erling Moxnes*

Institute of Information Science, University of Bergen

P.O.Box 7800, N-5020 Bergen, Norway

and Foundation for Research in Economics and Business Administration, SNF

1

Abstract

A frequent argument against efficiency standards is that they prohibit products that represent optimal choices for customers and thus lead to reduced customer utility. In this paper we propose and test a method to estimate such losses. Conjoint analysis is used to estimate utility functions for individuals that have recently bought a refrigerator. The utility functions are used to calculate the individuals' utility of all the refrigerators available in the market. Next the method is used to estimate losses due to energy efficiency standards that remove products from the market. The method is found to give accurate estimates that are consistent with other data on customer choices. Contrary to previous claims, we find that efficiency standards can lead to increased utility for the average customer. This is possible because customers do not make perfect choices in the first place.

Keywords: information, search behaviour, standards, conjoint, energy, utility

JEL Classification Code: Q49, Q39.

PsycINFO Classification Code: 3670

Revised and shortened version to appear in *Journal of Economic Psychology*

---

* Corresponding author. Tel.: +47 55584119 Fax.: +47 55584107

E-mail adress: Erling.Moxnes@ifi.uib.no

1. Introduction

In recent years the use of energy efficiency standards have increased. Nadel (2002) reports standards to be used in 17 countries plus the European Union and to "have been set on more than 35 products, with refrigerators, air conditioners, ballasts, and freezers being the most common." There seems to be general agreement that these standards lead to reduced energy consumption, oil imports, and CO_2-emissions, Greene (1998), Koomey et al. (1999), and Nadel (2002). According to the same authors, there also seems to be evidence to the fact that standards are cost effective, reductions in energy expenditures typically exceed the costs of improving the energy efficiency. However, this measure is likely to vary considerably between products. Koomey et al. estimate the US cost savings to be 165 times larger than the costs incurred by the US government to set the standards.

Still the use of standards is controversial. According to Sutherland (1996): "-there is no evidence that conservation programmes, such as energy efficiency standards for appliances, maximise net benefits as measured by consumer preferences." His arguments are that "consumers have different preferences, and -- consumer preferences cannot be determined by the state." It is not difficult to find products and standards for which this claim is self evident, e.g. mandating that everybody should wear shoes of a given size. Since the individuals know best how the shoe fits, a standard would reduce utility for most people. On the other hand, examples also exist to the contrary, e.g. who would go to a brain surgeon that does not satisfy a minimum standard with respect to competence? In this case the quality of the surgeon is all that matters, individual differences between brains and lack of knowledge about these differences are of little or no importance for the choice of standard. From these two examples it should be clear that the question about net benefits of standards is an empirical one.

The purpose of this paper is to suggest and test an empirical method to measure if average customer utility is increased or decreased by product standards. The case is refrigerators and the standard regulates energy efficiency. To introduce the method, first note one particular condition that must always be satisfied for a standard to yield an increase in customer utility. All customers cannot make perfect choices. If they did, there is only potential for reduced utility by setting a standard. However, if people make imperfect choices in the first place, it is no longer self-evident that a standard reduces utility. If for instance the standard removes both the optimal product and some bad choices for an individual, the expected outcome of a somewhat random search process could go either way.

A main reason to expect imperfect choices is that people tend to use highly simplified procedures when searching for durable products with many attributes. Since energy efficiency is often one of the less important attributes, it is paid limited attention in the search process. Previous studies indicate considerable losses due to non-optimal customer choices of energy consuming durables, see e.g. Fisher and Rothkopf (1989), Gately (1980), Ruderman et al. (1987), and Sutherland (1991). Some of these losses are claimed to be caused by systematically biased decisions, e.g. customers using discount rates that by far exceed interest rates on loans. In this study, however, we accept and use revealed customer preferences as they are. To minimise biases we make an effort to lessen the congnitive burden for the investigated customers.

In one special case the effect of a standard on customer utility can be found simply by considering data on energy efficiencies and related product costs. Assume a product where energy efficiency is an independent attribute that can be chosen freely for all the combinations of the other attributes of the product. Also assume that energy prices, discount rates, rates of utilisation, and values of waste energy do not differ among customers. In this case one can figure out the optimal efficiency of the product and set the standard thereafter. All customers would be better or equally well off. Problems with this method arise if some of the assumptions cannot be justified. We analyse a case in which customers have different preferences and where energy efficiency may be related to other product attributes. Then it is unclear *apriori* what the effect of a standard will be on customer utility.

We suggest the following method: A conjoint analysis is used to estimate individual utility functions for customers that have recently bought a refrigerator. Using these utility functions, we identify each customer's utility of all the refrigerators in the market. By comparing the estimated utility for the chosen product to the products that yield the maximum and minimum utilities, we find the relative utiltiy of the actual choice. A related method to measure losses is previously applied by Gupta and Ratchford (1992). We use a Monte Carlo analysis to test the accuracy of the method. Finally, we use the estimated utility functions to find customer losses (or gains) due to energy efficiency standards that remove refrigerators from the original choice set. We find that most efficiency standards lead to an increase in average customer utility.

Even though a standard has a positive effect on average customer utility, and it could be considered to be a public good, there is still an argument that is sometimes used to impede the implementation of standards. One may argue that any product is a potential candidate for standards, not only energy consuming equipment. Thus the question arises which standards should be prioritised. The general answer is that all welfare enhancing public goods should be provided. If forced to prioritise, energy consuming products should get high priority because of externalities such as climate change, local pollution, and vulnerability to supply shortages. We recognise that in this case the first best policy is to increase prices of energy by use of taxes or quota limitations. However, there seems to be considerable political resistance towards taxing energy appropriately. This is why externalities provide another argument for the use of standards, as a second best policy. There is also a dynamic aspect of this argument. By using standards to reduce the dependence on energy, it will be easier to raise energy taxes to secure a first best policy in the future. In this paper we will base the reference case on current energy prices, however, we will also find the effect of standards in case the energy price is doubled.

2. Method

First we define optimal choices and describe how losses from non-optimal choices can be estimated. Then we describe the conjoint technique which is used to estimate individual utility functions. We present a Monte Carlo test of the method, and finally we describe how to estimate losses or gains due to efficiency standards.

2.1. Definition of losses from non-optimal choice

Search costs, imperfect product information, and imperfect information treatment can cause customers to end up with non-optimal products. To measure an individual customer's loss, one could use compensating variation, c

$$u(p^*,\mathbf{q}^*) = u^* = u(p'-c,\mathbf{q}') \tag{1}$$

where $u(p^*,\mathbf{q}^*)$ denotes the utility of the optimal choice for the customer (the product with the highest utility), where the optimal product is characterised by a price p^* and a set of other attributes $\mathbf{q}^*=\{q_1^*,...,q_J^*\}$. The customer's actual choice of product is characterised by a price p' and a set of attributes $\mathbf{q}'=\{q_1',...,q_J'\}$. The compensating variation c is the necessary price decrease to achieve the same utility from the actual choice as from the optimal choice. Thus c represent a measure of customer loss. If the utility function $u(p,\mathbf{q})$ can be estimated and the choice of product $(p',\mathbf{q}')$ is known, the value of c can be estimated.

In the ensuing section we will explain how the conjoint technique is used to estimate linear utility functions $u(\alpha)= a_p p+\alpha_1 q_1+\alpha_2 q_2+...+ \alpha_J q_J$, where $\alpha=\{\alpha_p, \alpha_1,...,\alpha_J\}$ is a vector of estimated attribute weights. Using such a linear utility function we can solve Equation 1 for the compensating variation

$$c=(u^*(\alpha)-u'(\alpha))/(-\alpha_p) \tag{2}$$

where $u^*(\alpha)$ is the estimated utility of the optimal choice, $u'(\alpha)$ is the estimated utility of the actual choice and α_p is the estimated weight on the price attribute.

There is one important practical problem with the use of Equation 2, it is very sensitive to estimation errors in α_p. If α_p unexpectedly comes out with a positive sign, the compensating variation will become negative. If α_p is close to zero, the compensating variation will become very large. For this reason we choose another measure of loss which is much more robust. We term this measure *relative utility*

$$r = \frac{u'(\boldsymbol{\alpha}) - u^m(\boldsymbol{\alpha})}{u^*(\boldsymbol{\alpha}) - u^m(\boldsymbol{\alpha})} \tag{3}$$

where $u^*(\alpha)$ and $u'(\alpha)$ denote respectively the utility of the optimal and the actual choice, while $u^m(\alpha)$ denotes the utility of the product with the lowest utility for the individual. Thus to calculate the value of r one must identify the products that yield the optimal and the minimum utility. This is done by enumeration. As long as utilities differ among products the denominator will be positive, and it will not be highly sensitive to estimation error. When the actual choice is equal to the optimal choice, r equals 1.0. When the actual choice is the product with the minimum utility, r equals 0.0. The weakness of this measure compared to compensating variation is that we get an index rather than a measure in currency units. Since our major concern is to find out if utility goes up or down due to efficiency standards, this is not a major concern. Using a relative utility measure, we will not be able to compare eventual benefits of of standards with the costs of implementing them, however this seems to represent no major loss in usefulness.

Finally, when using data for many individuals, we are interested in aggregate losses. Taking the average of compensating variations c is a rough measure that can be criticised. A simple average puts the same weight on all individuals not recognising that a given loss may be perceived differently among individuals with different marginal utility of money. Taking the average R of individual relative utilities r also means that equal weights are applied. Thus, this average can also be criticised as a measure of aggregate loss. However, as for compensating variation, we see no simple way to improve the aggregate measure. Since relative utility is bounded on a scale from zero to one for each individual, no individual can have a dramatic effect on the average. Using compensating variation, one can not be assured of this since estimation errors could lead to large individual variations in c.

Since the measure of relative utility is not the standard one, it is reassuring that the results obtained here are almost identical to the results obtained in Hatlebakk and Moxnes (1996) using compensating variation and a number of auxilliary assumptions.

2.2. The conjoint technique

A conjoint technique is used to estimate individual utility functions, Green and Srinivasan (1990). The functional form of the utility function and the product attributes are determined *apriori*, while the technique is used to estimate the parameters of the utility function. According to Green and Srinivasan (1978): "The linear-compensatory model .... can approximate the outcomes of other kinds of decision rules quite closely". With this advice, and in order to save degrees of freedom, we apply a linear additive functional form with parameters or attribute coefficients a_j, $j = 0, p, 1, \dots, J$:

$$u(p,\boldsymbol{q}) = a_0 + a_p\, p_k + a_1\, q_1 + \dots + a_J\, q_J \qquad (4)$$

We use a full profile conjoint design where respondents are asked to rank order a set of products, according to their probability of buying them. The products are described by a set of common attributes which differ in attribute levels. Rank numbers ρ_k for product k, are regressed against attribute values for each respondent

$$\rho_k = a_0 + a_p p_k + a_1 q_{1k} + \ldots + a_J q_{Jk} + \varepsilon_k \qquad (5)$$

where ε_k is the random error. In accordance with current practice in conjoint analysis, OLS is used to estimate the α vector.

At the outset it may seem overly optimistic to judge the utility of actual purchases by an utility function found by the conjoint technique. The motivation for doing this is one important difference between actual purchasing strategies and the conjoint technique. Actual purchasing involves costly search for information about the attributes of different models. Investigations show that customers tend to investigate only a few models and to be less than fully informed about their attribute values (see Table 6 for some numbers and references). The conjoint technique presents the customers with well organised and precise information about all the models that are to be ranked. Thus using the conjoint technique, customers are not constrained by the search problem, and they are stimulated by the organisation of attribute information on cards to apply efficient heuristics to rank the alternatives.

This is not to say that the conjoint technique produces perfect estimates of utility functions. We rely on assumptions about a linear additive functional form, about a given set of attributes, and about the ability of respondents to rank order the products without a common systematic bias.

2.3. A Monte Carlo test of relative utility estimation

The standard errors of the conjoint estimates of the α vector reflect model misspecifications and lack of precision in customer rankings of cards. The greater these errors, the greater the chance that we will report values of r that reflect errors of the method rather than of the customer choices. The error has two components, a certain spread measured by a standard deviation and a systematic bias caused by non-linearities. A Monte Carlo analysis quantifies these errors.

From Equations 3 and 4 one can see that uncertainty in α, except the constant term, implies uncertainty in r. The potential bias in r is harder to recognise. To illustrate, assume that an individual has made a perfect choice such that the true loss is zero, r=1.0. In this case an erroneous estimate of α can lead us to identify a wrong product as the one giving the optimal utility. Consequently we may estimate a value of r which is less than the true value of 1.0. Since we will never estimate a value of r greater than 1.0, on average we should expect a downward bias in this case. Similarly, if a customer has made the worst possible choice, identifying the wrong minimum choice could lead to an upward bias in the estimate of r. For a set of customers that have made a mix of good and bad choices, the bias in the average relative utility R could go in either direction. It is possible to improve the method to reduce the bias[1], however this would complicate the presentation of the method, the programming and the Monte Carlo test.

Using the Monte Carlo technique, we first establish a benchmark for each individual's relative utility by using Equation 3 under the assumption that the estimated parameter vector α is the correct one. We make two different assumptions about simulated customer choices. First we assume that the customer makes an optimal choice, in which case the benchmark for the relative utility $\bar{r}_m$ will be 1.0 for all Monte Carlo replications, m=1, 2, ... M, where M=100. Second we assume that the customer chooses at random with equal probabilities of choosing all the available products. In this case the benchmark for the relative utility $\bar{r}_m$ will reflect the average utility of all available products. Since real customers are likely to do better than the random strategy and worse than the optimal choice, real actual choices are likely to be somewhere between our two assumptions.

Then we turn to the Monte Carlo estimates of relative utility based on erroneous parameter estimates. The parameter estimates

---

1 A more accurate method would have been to set the utility of the optimal product equal to the utility of the chosen product when there is no statistically significant difference between u' and u*, yielding r=1.0 in these cases. In this case the bias would be reduced. The remaining bias would be a function of the choice of significance level.

$$\hat{\alpha}_{mj} \sim \text{Triangular}(\alpha_j, \sigma_j) \tag{6}$$

for attribute j and Monte Carlo replication m, are drawn from a triangular distribution with mean α_j and standard deviations σ_j, where α_j and σ_j follow from the conjoint analysis. The estimated relative utility

$$\hat{r}_m = \frac{u'(\hat{\boldsymbol{\alpha}}_m) - u^m(\hat{\boldsymbol{\alpha}}_m)}{u^*(\hat{\boldsymbol{\alpha}}_m) - u^m(\hat{\boldsymbol{\alpha}}_m)} \tag{7}$$

for one Monte Carlo replication m is found by a variant of Equation 3 where the correct parameter vector $\boldsymbol{\alpha}$ is substituted by the erroneous estimate $\hat{\boldsymbol{\alpha}}_m$. As in Equation 3, it is necessary to calculate the utility of all products to identify the products yielding optimal and minimum utility. The difference between the benchmark $\bar{r}_m$ and the estimate $\hat{r}_m$ defines the bias

$$b_m = \bar{r}_m - \hat{r}_m \tag{8}$$

The expected bias μ_b for each respondent is found by taking the average of b_m over all Monte Carlo replications M. To get a measure of the accuracy of the method we calculate the standard deviation σ_b of b_m over all M replications.

Finally the above procedure is repeated for all the N respondents participating in the study. By taking averages over individuals we obtain a more representative estimates of the bias and the accuracy. $\bar{R}$ denotes the average of individual benchmarks $\bar{r}$, which in turn are found as averages of $\bar{r}_m$ for each respondent. $\hat{R}$ denotes the average of individual estimates $\hat{r}$, which in turn are found as averages of $\hat{r}_m$ for each respondent. B denotes the average of the expected individual biases μ_m and Σ denotes the average of the individual standard deviations σ_b.

2.4. Calculation of losses or gains due to efficiency standards

An energy efficiency standard reduces the choice set. Only refrigerators that have a yearly energy consumption less than the standard, $q_e<q_S$, will be allowed in the market. Again Equation 3 can be used to calculate the relative utility for each individual

$$r_S = \frac{u'_S(\boldsymbol{\alpha}) - u^m(\boldsymbol{\alpha})}{u^*(\boldsymbol{\alpha}) - u^m(\boldsymbol{\alpha})} \tag{9}$$

where $u^m(\boldsymbol{\alpha})$ and $u^*(\boldsymbol{\alpha})$ are given by the original choice set. Thus the utility of the choice after the standard has been set $u'_S(\boldsymbol{\alpha})$ is seen in light of the original range of utilities, i.e. the range from optimal to minimum utility before any products have been removed by the standard. We define the loss

$$l = r - r_S = \frac{u'(\boldsymbol{\alpha}) - u'_S(\boldsymbol{\alpha})}{u^*(\boldsymbol{\alpha}) - u^m(\boldsymbol{\alpha})} \tag{10}$$

as the difference between r and r_S. Because the denominator is the same for r and r_S in Equations 3 and 9, we see from Equation 10 that the loss is directly related to the difference between the utility of the actual choice $u'(\boldsymbol{\alpha})$ and the choice made after the standard is put in place $u'_S(\boldsymbol{\alpha})$. L denotes the average of individual losses. Finally, note that a possible bias in the estimate of the relative utility for the original choice carries over to the relative utility of the choice with a standard. Important to note, these two possible biases cancel out when r_S is subtracted from r to calculate the loss l due to the standard.

There is one challenge in calculating r_S, an assumption has to be made about how customers choose given the new restricted choice set. First note that customers typically buy refrigerators with very long time intervals, it is not such that they make a choice the day before the standard is implemented and then they revisit the same store the day after to make another choice. Thus when customers replace their refrigerators they will typically face a totally new collection of refrigerators and they will not be influenced by the data they gathered the last time they bought one. Thus we cannot assume that they will repeat their previous choice even if that choice is allowed by the standard. We can only speculate that they will use the same simplified search procedures that they used last time. This suggests that the expected quality of the choice they make will be the same. However, using even optimal search procedures there is a considerable random element. This implies that we cannot make reliable predictions of individual choices. We can only choose a procedure that maintains the average goodness of the choices made by all individuals. A simple, and probably best possible way to do this is to assume that each individual will make a choice that gives the same (closest possible) relative utility before and after the standard. Note that for this comparison the relative utility after the standard

$$r_S^S = \frac{u'_S(\boldsymbol{\alpha}) - u^m{}_S(\boldsymbol{\alpha})}{u^*{}_S(\boldsymbol{\alpha}) - u^m{}_S(\boldsymbol{\alpha})} = r \tag{11}$$

must be based on the choice set allowed by the standard, with optimal utility $u^*{}_S(\boldsymbol{\alpha})$ and minimum utility $u^m{}_S(\boldsymbol{\alpha})$ based on the available products only. The new choice is found by enumeration. The product that brings r_S^S and r closest to each other is chosen. Over a large sample of respondents, minor errors, due to a lack of products that ensure perfect equality, should be expected to cancel out. Hence this procedure should ensure that average estimated effects of standards are not explained by changes in the quality of the decisions. The expected effect on average utility is only due to the standard.

3. Design of the investigation

In order to estimate losses from non-optimal choices, we need to know both actual choices and the choice set available at the time of purchase. To limit the investigation to one choice set, and to be up to date, we investigated the population of customers that had bought a refrigerator recently. We got lists of customers who had recently bought a refrigerator from the "Bonus" chain, which has four stores in the district of Bergen, and which is the largest seller of refrigerators in the district. 180 customers, who bought a refrigerator in the period April 26th to June 24th 1995, got a letter with the conjoint cards. The customers were asked to rank order the cards, to write their rank ordering in an attached scheme, and to return the scheme in a pre addressed and stamped return envelope.

To stimulate participation, the customers were told that they had a chance of 1/100 to win 50 tickets in a state lottery with a purchasing price of NOK 1000 (Euro or USD 120). To take part in the lottery, respondents had to write their name and address on the mail back envelope, which all did. They were told that the mail back envelope would be separated from the survey scheme. About a week later a female student assistant phoned them, and asked if they needed some help, and also asked them to return the survey scheme. As a final strategy, she offered one or two tickets in the state lottery if they would return the scheme. Two weeks later a reminder was sent to the respondents who had agreed to return the survey scheme.

To increase reliability of the estimated parameters in a conjoint analysis, attribute levels should have a large range, and no correlations between attribute levels should exist. However, to ensure validity, the range should be within the actual range of attributes in existing products. This trade-off lead us to perform three separate studies of respectively small, medium and large refrigerators. Thus, the 180 customers were divided into three samples of 81, 44 and 55 customers, who bought respectively small (less than 200 litres), medium (200-249 litres) and large refrigerators (250 litres and above). From these samples we got 17+14+15=46 useful answers, which imply a response rate of 25.6 percent. While this is a fairly low response rate by most standards, we will see that the sample is large enough to get good indications about the accuracy of the method and to encourage further research with more effort devoted to increasing participation.

Unfortunately, eleven of the respondents who were registered as buyers of small refrigerators in the received name lists, turned out to have bought large refrigerators. Hence they got the wrong set of cards. They had all bought the same refrigerator with interior volume of 251 litres. Since 251 litres is almost inside the range 200-249 litres of the medium refrigerators, we grouped the buyers of this product together with the buyers of medium refrigerators when we test the accuracy of the method and when we investigate efficiency standards. Thus we end up with respectively 7+24+15=46 respondents in the respective size classes.

The following attributes were included. Energy efficiency follows from the focus of the study. Attributes which are likely to be negatively correlated with energy efficiency were included to ensure that energy efficiency is not provided without costs. Price is an obvious candidate. In addition extra insulation either decreases interior volume or increases exterior volume. Since width and depth are standardised, exterior volume is measured by the height of the refrigerator. Thus we include inside volume and height in the design. To increase validity, the most important attributes should be included, because the remaining parameters can be biased if important attributes are excluded. After discussions with sales persons and some customers, we included quality as a fifth attribute in the conjoint task. If quality had not been included, some customers would probably use price as a proxy for quality. In that case, we could not have interpreted price as a pure cost variable. We divided the quality attribute into three different classes, low, medium and high quality. The models were classified into these groups by one sales manager in a Bonus store. A sales person in another (not Bonus) store (almost) agreed with this ranking. In the introduction to the conjoint task we described quality as including noise and durability. Furthermore, we use "a brand produced in Germany such as Siemens" as an example of high quality, and "a brand produced in Russia such as Atlas" as an example of low quality. Attributes such as colour and equipment, with small or no variation between products, were not included. The discussions with customers were organised according to the rank ordering method, found to give good results with little bias in Breivik and Supphellen (2003).

To be able to identify the optimal choice for each individual, we made efforts to avoid that individuals misinterpreted the attribute values in the conjoint design. To compare energy costs and prices, respondents have to estimate the present value of future energy costs, utilising their real market interest rate. Such calculations are difficult, and one cannot expect correct results. We tried to reduce these problems by using energy costs per year as the attribute for energy efficiency, and by reminding respondents that the chosen refrigerator will use energy during its entire lifetime, that interest costs imply decreasing value of money saved on energy, and that total costs are reduced if the refrigerator diminish the need for space heating.

In Table 1 we have listed the range of all attributes except quality for the real choice set and the attribute values for the products described on the conjoint cards. Energy costs per year were based on an electricity price of NOK 0.4/kWh (Euro or USD 0.05/kWh). In the experimental design the same attribute differences were used for all size classes. There exist 3^5=243 different combinations of the attribute levels in the conjoint design. From these we have chosen an orthogonal design (no correlations between attributes) of 25 cards, using a commercial conjoint design algorithm.

Table 1. Real and experimental attribute values.

Attributes	Real range	Conjoint range
Small refrigerators		
Inside volume (litres)	109-197	110-150-190
Height (cm)	85-132	85-100-115
Energy cost (NOK)	43.8-189.8	43.8-131.4-219
Price (NOK)	1690-4890	1500-2250-3000
Medium refrigerators		
Inside volume (litres)	197-251	190-230-270
Height (cm)	108.5-138.5	115-130-145
Energy cost (NOK)	43.8-219	43.8-131.4-219
Price (NOK)	1990-4890	3000-3750-4500
Large refrigerators		
Inside volume (litres)	245-374	270-310-350
Height (cm)	125-187	145-160-175
Energy cost (NOK)	51.1-153.3	43.8-131.4-219
Price (NOK)	3590-6190	4500-5250-6000

The 46 respondents bought 4+6+8=18 different models in respectively the small, medium and large size group. Nine different brands were represented. In the ensuing analysis of standards we also included models from Gram, which is a high quality brand not sold by the Bonus stores. Two of these models were included both as small and medium models (volume of 197 litres). One of the Gram models is a low energy refrigerator. We also included a low energy model from Electrolux.

Finally, the respondents receiving the conjoint cards were also asked which attributes they considered most important.

4. Estimated attribute weights

From the rank orderings of products we estimate the α-parameters for each respondent. The quality attribute is modelled by two dummy variables, with medium quality as default. The other attributes are quantitative variables. Table 2 summarises the results. When not otherwise stated, the table refers to significant parameters at the 10 percent level.

As many as 37 of the 46 respondents had R^2 values above 0.79, indicating a high frequency of consistent rankings. The average R^2 was 0.80, while the average R^2 for respondents with $R^2>0.79$ was 0.94. Counting significant parameters (row 1) low quality (36) comes out as the most important attribute, followed by high quality (31), energy costs (23), volume (19), price (17) and height (14). Most of the significant parameters have the expected signs (row 2).

Hence, all averages of the significant parameters have the expected sign (row 3). The standard deviations of the average significant parameters (row 4) shows that all averages are highly significant with the exception of height.

Table 2. Results from the OLS estimation. Significant parameters when not otherwise stated (*N*=46)

	Const.	Height	Volume	Low quality	High quality	Energy costs	Price	R^2
1 No. of sign. parameters	34	14	19	36	31	23	17	37*
2 of which with expected sign		10	16	36	31	22	16	0.94#
3 Average	28	-0.15	0.060	-7.9	6.0	-0.025	-0.0036	0.80
4 St. deviation of average	5.8	0.092	0.015	0.55	0.45	0.0065	0.0009	0.04
5 Average WTP		-624	672	-2207	1680	-621	-750	
6 Average, all parameters	22	-0.037	0.023	-6.2	4.7	-0.014	-0.0014	
7 St. dev., average all param.	4.7	0.031	0.0081	0.65	0.45	0.0038	0.0004	0.04
8 Average WTP, all parameters		-397	676	-4511	3438	-873	-750	
9 Average, small refrigerators	16	-0.11	0.087	-8.1	4.8	-0.024	-0.0042	0.92
10 Average, med. refrigerators	44	-0.26	0.024	-7.4	6.2	-0.032	-0.0043	0.72
11 Average, large refrigerators	36	-0.11	0.023	-8.0	7.8	-0.022	-0.0019	0.76
12 Sign. difference, small-med.?	Yes	No	No	No	No	No	No	Yes
13 Sign. difference, small-large?	No	No	Yes	No	Yes	No	No	Yes
14 Sign. difference, med.-large?	No	No	No	No	Yes	No	No	No
15 Average, reminded respond.	26	-0.09	0.041	-8.9	6.9	-0.017	-0.0036	0.68
16 Average, other respondents	29	-0.19	0.069	-7.1	5.5	-0.030	-0.0035	0.92
17 Sign. difference?	No	No	No	Yes	Yes	No	No	Yes

* Number of subjects with $R^2>0.79$

\# Average R^2 for respondents with $R^2>0.79$.

Row 6 shows averages taken over all 46 respondents including significant and insignificant parameters. Similar to the case with significant parameters, standard deviations for the averages (row 7) show that all averages except the one for height, are significantly different from zero.

When the average parameter values are multiplied with the standard attribute differences used in the conjoint design (Table 1), we get a measure of attribute importance. In Table 2 this measure of importance is divided by the average parameter for price to get a willingness to pay (WTP) measure in NOK. Significant parameters (row 5) and all parameters (row 8) give nearly the same rankings: it is most important to avoid low quality (compared to medium quality), followed by the preference for high quality (compared to medium quality).

There are some not very important differences between the three groups (rows 9-14). 22 of the 46 who responded, got a telephone reminder. We have compared parameters and R^2 values of this group to the results of those who answered without a reminder (rows 15-17).

The early responders produced the most consistent rankings having significantly higher R^2 values than the laggards. Significant differences between average parameter values are less than 25 percent of average estimates.

Then we compare the self-reported "most important attributes" to WTP-estimates for each individual (row 8). In the self-reports each individual is free to name as many attributes as he or she wants. For each individual we select WTP-estimates for the same number of attributes, where the selected attributes are those with the highest WTP-scores (for consistency only one quality attribute is counted). Table 3 shows that the frequencies found by the two methods are quite similar. Both methods indicate that we have included only important attributes.

Table 3: Self-reported attribute importance versus WTP-estimates, percentages.

	Height	Volume	Low quality	High quality	Energy cost	Price	Average
Frequency of self-reported importance	37	72	89		54	76	66
Frequency of high estimated WTP	63	41	85	89	61	52	65
Frequency of matching predictions	70	65	83	78	80	67	74

Finally, we see how frequent self-reported importance coincides with incidents of high estimates of WTP. A match is reported if both methods signal importance or if both methods signal no importance. The third row in Table 3 shows that the match is clearly higher than 55 percent, which is the match that would follow if all estimates were random[2].

5. Monte Carlo test of the relative utility estimation

Table 4 shows the Monte Carlo results for each of the three groups. In case customers are assumed to make optimal choices, the benchmark for relative utility $\overline{R}$ is 1.0. The average Monte Carlo estimates for relative utility $\hat{R}$ are not much lower than the benchmark values. The average bias is less than 0.06 in all size classes. The accuracy of the predictions for the individuals is also good. The average standard deviation Σ is less than 0.10 in all size classes. The accuracy of the prediction of relative utility for a group average is of course even better (divide the standard deviation by the square root of group members n).

2 This benchmark is higher than 50 percent because there is a higher probability of importance than no importance for both methods: (1-0.65)*(1-0.66)+0.65*0.66=0.55

The subjects with low R^2-values tend to increase the bias and to reduce the accuracy. This is seen if we compare the results for all respondents with the results for subjects with $R^2>0.79$ (five subjects are removed from the medium size group and four from the large size group). Now the largest average bias is 0.04 and the largest average standard deviation is 0.06. The 9 excluded subjects, who all have R^2-values below 0.48, explain roughly half the bias and half the standard deviation of the results when all individuals are included.

Next we consider the relative utility in case customers are assumed to choose randomly among the different products. First we note that the average benchmark $\overline{R}$ varies from 0.48 to 0.61 when all subjects are included. The average bias B is close to zero in all size groups. All average standard deviations are lower than 0.14. If we again remove subjects with R^2-values below 0.79, biases equal 0.00 and the accuracy improves with the largest average standard deviation being equal to 0.11.

Table 4: Relative utilities, biases in relative utility, and standard deviations of biases in relative utility.

Size group refrigerators	n	Average benchmark $\overline{R}$	Average estimate $\hat{R}$	Average bias B	Average st.dev. of ind. biases Σ
Benchmark based on optimal decisions					
Average over all respondents					
Small	7	1.00	1.00	0.00	0.01
Medium	24	1.00	0.95	0.05	0.06
Large	15	1.00	0.94	0.06	0.10
Average over respondents with $R^2>0.79$					
Small	7	1.00	1.00	0.00	0.01
Medium	19	1.00	0.99	0.01	0.02
Large	11	1.00	0.96	0.04	0.06
Benchmark based on random decisions					
Average over all respondents					
Small	7	0.56	0.56	0.00	0.04
Medium	24	0.61	0.60	0.01	0.10
Large	15	0.48	0.49	-0.01	0.14
Average over respondents with $R^2>0.79$					
Small	7	0.56	0.56	0.00	0.04
Medium	19	0.63	0.63	0.00	0.05
Large	11	0.49	0.49	0.00	0.11

Due to a limited number of Monte Carlo replications average relative utilities may deviate by around 0.01.

We conclude that in our case the method produces only a modest bias, even when customers make optimal choices. Furthermore the method is quite accurate in predicting individual relative utilities. Average relative utilities for groups of subjects are even more accurate. As always when using models, we cannot rule out biases and inaccuracies caused by model errors. However, the high R^2-values suggest that it will be hard to discriminate between the current linear utility function and more elaborate models since there is only a minor potential for improvement.

6. Estimates of the relative utility of actual choices

Table 5 shows the estimates of average relative utilities *R* for the three size classes, using Equation 3. The second line in the table repeats the average relative utilities reported in Table 4 from the Monte Carlo simulation with random choices. Except for the group with small refrigerators, actual choices produce higher relative utilities than what would result from a random choice. However, the relative utilities are quite a distance away from the maximum relative utility of 1.0.

To see if low utilities are explained by poor estimates of utility functions we remove all respondents with a R^2-value less than 0.79. This has an effect in the case with large refrigerators, for small and medium refrigerators the effect is negligible. This means that also in the case where respondents rank order the conjoint cards systematically according to attribute values, relative utilities are far from optimal. This lack of perfection in real customer choices may seem surprising, however, it seems consistent with the following evidence.

Table 5: Estimated relative utility compared to results of Monte Carlo simulations with random choice and to estimates based on a given search strategy.

	Small	Medium	Large	All groups
All respondents				
Estimated average relative utility of actual choice	0.47	0.67	0.59	0.62
Monte Carlo av. relative utility when random choice	0.56	0.61	0.48	0.55
Respondents with R^2>0.79				
Estimated average relative utility of actual choice	0.47	0.68	0.73	0.65
Monte Carlo av. relative utility when random choice	0.56	0.63	0.49	0.57
Estimates based on search strategies/uniform distr.				
Empirically based number of searches	0.64	0.64	0.64	0.64
Random choice	0.50	0.50	0.50	0.50

A first indication of non-optimal choices is provided by customers who want to change newly bought refrigerators. Own experiences as well as reports by the retailers we were in contact with indicate that people at times want to change a newly bought product. Most probably, more customers than those that actually take an action are dissatisfied. Actions can be hampered by uncertainty about own preferences, fear of making another bad choice, considerations of sunk costs (Kogut, 1990), tendencies to postpone actions (Akerlof, 1991), and defensive actions to cover poor choices in the first place (Argyris, 1990).

Second, an investigation by Gately (1980) shows that customers miss out on dominating alternatives (equal to or better on all attributes). This happens even when the products are placed next to each other in the store.

Third, investigations of purchasing behaviour show that many customers investigate only a small number of options, see Table 6. In particular, note the high fraction of customers that make only one search. If there is no information search beyond what is reported, more than 40 percent of the customers choose at random. Probably, customers often possess some *apriori* information, and they are likely to get useful advice from salespersons. However, the fewer products they consider, the more likely it is that they come to ignore important attributes.

Table 6: Empirical estimates of number of searches and of the percentage with one search only.

	Average no. searches		% with only one search
	Dealers	Brands	
Newman and Staelin (1972), automobiles and household appliances	2.3	-	49
Dommermuth (1965), refrigerators	-	2.3	41
Claxton et al. (1974), household appliances	2.4	-	-
Brottemsmo and Moxnes (1991), automobiles	2.1	2.7	43

Fourth, normative search theory predicts non-optimal choices when there are search and information processing costs, see Hey (1981) who finds similar losses with optimal search strategies and simple, practical strategies. To quantify, we use a simple search strategy: investigate k products and choose the one with the highest utility. We assume that the products are uniformly distributed with a utility ranging from 0.0 to 1.0, i.e. the same range as the previous measure of relative utility. Then the expected relative utility of the search is simply given by $k/(k+1)$.[3] When only one product is searched, the relative utility will be 0.5.

3 The probability that utility is greater than some value u is given by the joint probability that all searches result in a utility below u, $F(u)=u^k$, and $dF/du=f(u)=ku^{k-1}$. Then we find the expected value of $f(u)$ on the range from 0.0 to 1.0 and get $k/(k+1)$.

To get a representative measure of the relative utility obtained by this search procedure, we use the data on the number of searches by Dommermuth (1965) in Table 6. The average, expected relative utility for the population is

$$R_{Search} = \sum_{k=1}^{7} w_k \frac{k}{k+1} \tag{12}$$

where w_k denotes the fraction of the population that make k searches. Fourtyone percent search only once, i.e. w_1=0.41. The other weights are set to obtain the average number of searches, 2.3, with a reasonably smooth distribution.[4] With these assumptions we find a relative utility of 0.64, shown for comparison in Table 5. This is very close to the estimated average relative utilities over all groups. The combined product seaches yield a relative utility which is 0.14 above the relative utility in the case of random search. This difference is larger than the corresponding ones found by Monte Carlo simulations for the small and medium sizes, it is about the same difference as the one found for large refrigerators. We conclude that the relative utilities based on conjoint estimates do not seem to be at odds with indications from other sources of information.

7. The effect of energy efficiency standards

Standards for refrigerators used in the U.S are set as linear functions of interior volume, Turiel et al. (1993). Here we simplify and assume one and the same efficiency standard for all size classes. Thus only refrigerators with an electricity consumption of less than q_S kWh/year are allowed to enter the market. This implies that we may underestimate the potential for standards somewhat, since we have only one rather than two parameters to adjust when searching for optimal standards.

Figure 1 shows average losses L (or gains if losses are negative) as a function of the efficiency standard, q_S. First consider the average loss over all size classes. The standard starts to rule out the refrigerators with the highest electricity consumption at 547 kWh/year. Then for the entire range of standards from 547 to 209 kWh/year, the loss is negative. In

---

4 The weights w_k for searches k from 1 to 7 are: 0.41, 0.26, 0.14, 0.09, 0.06, 0.03, and 0.01. The result is not very sensitive to likely variations in the distribution.

other words, the average customer gains in terms of utility. For standards below 209 kWh/year there is a loss, increasing to 0.073 relative utility units at the lowest possible standard of 128 kWh/year. This standard leaves only one refrigerator in each size class (this is also the situation at the lowest standard in the figure at 150 kWh/year). To put this loss in perspective, it is considerably less than the average loss caused by non-optimal choices in the first place, 0.35 relative utility units.

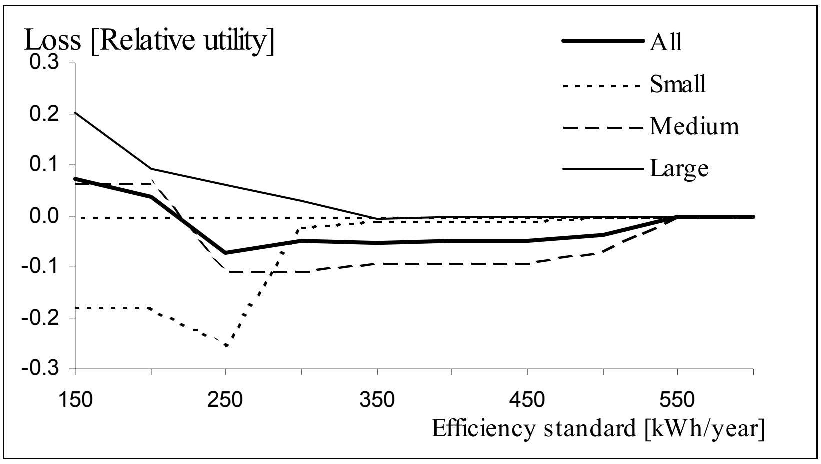


Figure 1: Average losses as functions of the efficiency standard

For medium size refrigerators the pictures is similar to the overall average. For large refrigerators there is a loss in utility when the standards go below 329 kWh/year, and it increases to 0.20 relative utility units towards the toughest standard. For small refrigerators the loss is always negative, being -0.18 relative utility units at the toughest standard. Thus, these results may suggest that the standard for electricity consumption should increase with the size of the refrigerators.

Next we want to see how sensitive the above results are to low R^2-values. The 9 respondents with $R^2<0.79$ are excluded. At the same time we also set all insignificant coefficients from the conjoint analysis equal to zero. This implies that one more respondent must be excluded since the remaining coefficients for this respondent give no variation in utility among products. The result is shown in Figure 2.

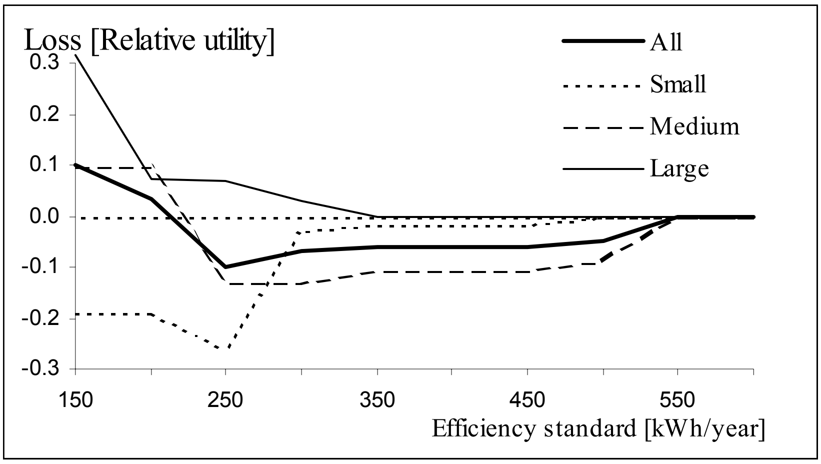


Figure 2: Average losses as functions of the efficiency standard (only respondents with R^2>0.79 and only significant conjoint estimates at the 10 percent level)

Compared to Figure 1, the result is hardly changed by this sensitivity test. Performing separate tests excluding only respondents with low R^2-values or only insignificant coefficients lead to similar results. Hence the insensitivity shown by Figure 2 is not caused by the two effects counteracting each other. Below we present figures for the effect of standards on the average attribute values. These figures are all based on the case with all respondents and all estimated coefficients. Also these results are insensitive to this assumption.

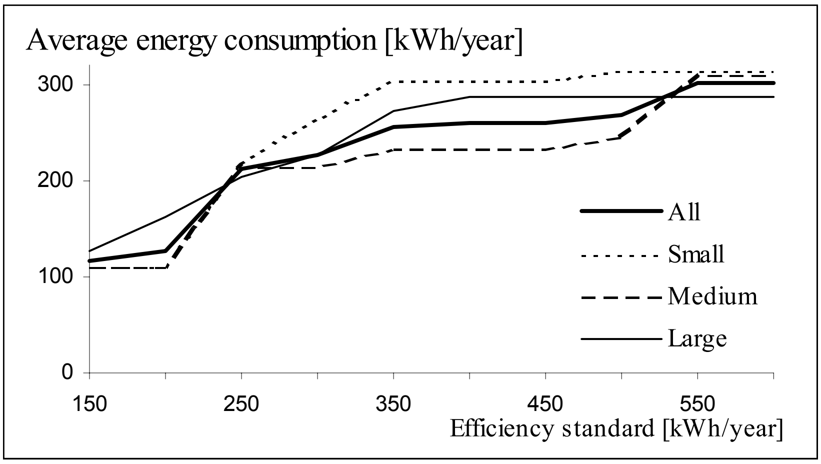


Figure 3: Average energy consumption per year as functions of the efficiency standard.

Figure 3 shows how the average energy consumption per refrigerator develops with the standard. We first consider the average over all size groups. By going from no standard to the toughest standard, energy consumption is reduced by 62 percent. The systematic effect of an increasingly tough standard is consistent with the empirical investigations showing that standards do have a significant effect on energy consumption. A standard of 200 kWh/year yields an energy saving of 58 percent, at a cost of an increase in relative utility of 0.038

relative utility units. A standard of 209 kWh/year gives a saving of 43 percent at a gain of 0.046 utility units.

Looking at the separate size groups we see that the development is quite similar in all groups. This is related to the fact that the ranges for energy consumption are almost the same in all size groups, respectively 110 to 328, 110 to 548, and 128 to 383 kWh/year in the small, medium, and large group. This was the main reason for using the same standard in all size groups.

Then we consider the effect of the standard on the most important of the attributes, quality. To investigate this effect we construct a linear measure of quality, where low quality has the value -1.0, medium quality has the value 0.0, and high quality has the value 1.0. Figure 4 shows that the overall average quality increases as the efficiency standard is reduced. The reason must be that the general quality attribute is correlated with energy efficiency. This seems natural since energy efficiency could be seen as one of several elements contributing to the overall quality of refrigerators. Thus, the energy efficiency standard drives low quality models out of the market.

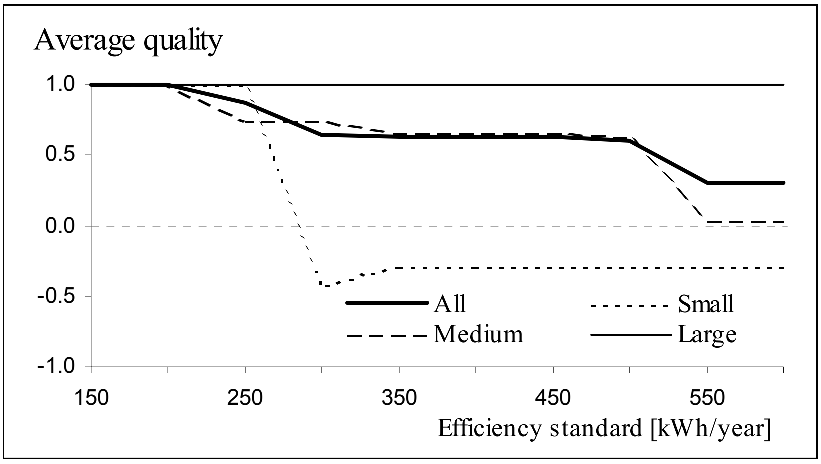


Figure 4: Average quality as a function of the efficiency standard.

Looking at the individual size classes we see that among large refrigerators there are only high quality models, for medium sized ones quality develops as for the overall average, while for small refrigerators there is a considerable jump in quality when the standard is reduced from 300 to 250 kWh/year.

What is the effect of the standard on the average interior volume of the refrigerators? Logically, more insulation should lead to less interior volume for a given height of the refrigerator. Figure 5 does not give a clear indication of such a reduction in volume. When looking at the individual size classes we see that at the toughest standard, the volume decreases somewhat for the medium class while in the small class it increases to 197 litres from 142 litres with no standard. Since volume has a positive value for most respondents, the increase in volume leads to increased utility.

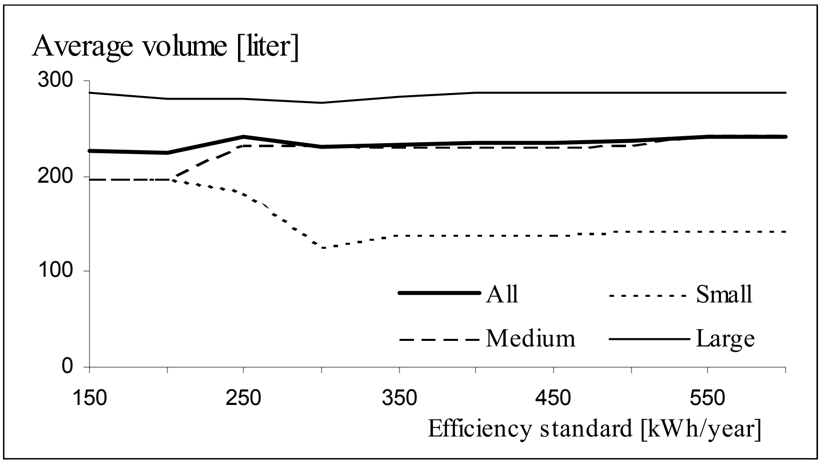


Figure 5: Average interior volume over all size classes as a function of the efficiency standard.

As quality and energy efficiency improves one should expect the price to rise. Figure 5 shows that this is what happens. It is interesting to note that there is hardly any increase in the average price of refrigerators as the standard is reduced from 500 to 250 kWh/year. Going from 250 to the toughest standard, the average price increases by 27 percent. From Figure 1 we see that this reduction in the standard turns a utility gain into a loss. The price increase is particularly strong for the small class. When this strong price increase does not lead to a similarly strong reduction in utility, it is because at the same time quality and interior volume both increase and energy consumption decreases.

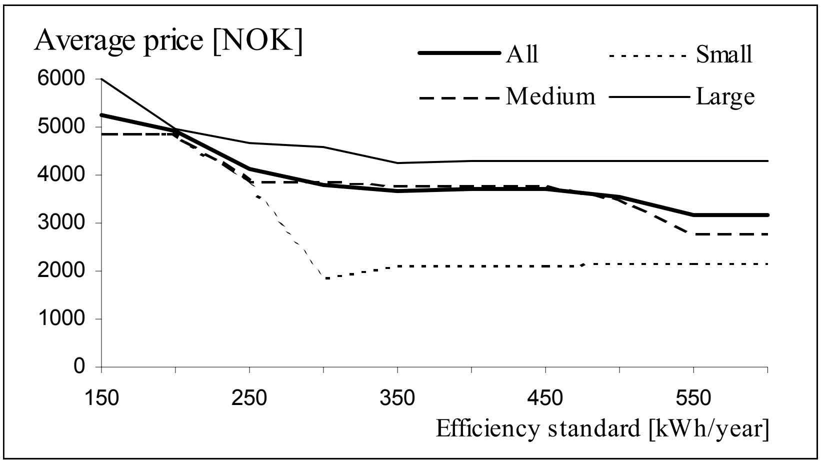


Figure 6: Average price over all size classes as a function of the efficiency standard.

Finally we test the sensitivity of our findings to the price of electricity? In the previous analysis we have assumed an electricity price of NOK 0.4/kWh (Euro and USD 0.05/kWh). This is around 50 percent of typical electricity prices to customers in most neighbouring countries. The major reason for this difference is that Norwegians have benefitted from cheap hydro electric power. Figure 7 shows the resulting losses in terms of relative utility when the electricity price is doubled. Compared to Figure 1 we see that losses are reduced as one should expect. A higher electricity price means that there is more to gain from efficiency improvements. The effect is not strong, at the most the average loss for all groups is reduced by around 0.023 relative utility units.

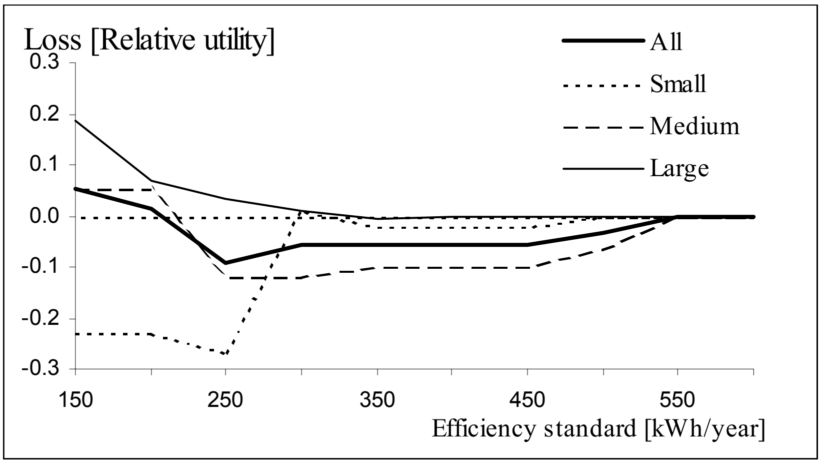


Figure 7: Average losses as functions of the efficiency standard, electricity price NOK 0.8/kWh.

7. Concluding discussion

We have proposed and tested a method to estimate customer losses due to energy efficiency standards. The main finding is that for most standards there is a gain in utility rather than a loss. Only for standards below 209 kWh/year there is a reduction in utility. When only the most efficient refrigerators are allowed in the market, average relative customer utility is reduced by 7 percentage points. These gains or moderate losses are possible because customers make imperfect choices in the first place, utility is on average around 35 percentage points below the maximum possible utility.

The above results are confined to a static context. The following dynamic factors should also be considered. First, standards should be announced ahead of the implementation time such that producers get a chance to adjust their models. Except for the fact that this seems fair to the producers, it will also help maintain sufficient competition in the market. Second, models that are produced in large quantities are cheaper per unit because of the effects of scale in production, advertisement, distribution, sales efforts and maintenance. Thus, an energy standard will lead to increased demand and reduced costs and prices of future low energy refrigerators. A simple test indicates that the price of the most efficient refrigerators must drop by 15 percent to prevent reductions in average utility at the toughest standard. Third, adaptations of other attributes than energy efficiency by producers could but does not have to contribute to further increases in utility given the toughest efficiency standard. Altogether this indicates that an efficiency standard of 128 kWh/year, giving a reduction in energy consumption of 63 percent could be both socially and privately preferable in the near future.

Further studies should aim for higher response rates, and should consider distributional effects. For instance what is the effect of standards on second-hand refrigerators, yielding a combination of low energy consumption and a low purchasing price. It may also be interesting to consider efficiency standards that depend on some other attribute, e.g. the volume of refrigerators, and to consider standards for other attributes than energy, e.g. certain aspects of quality. Finally, it is important to study other products for which energy efficiency standards seem appropriate. The question about the effect of standards on utility is an empirical one, and one should be careful not to generalise the quantitative results shown here.

Acknowledgements

A special thank to Magnus Hatlebakk, University of Bergen and SNF, who participated in the original project and who has made valueable contributions. I am also thankful for constructive comments from the editor and two anonymous referees. Thanks to the Bonus chain for helping out with data and namelists. The project has been financed in part by the Norwegian Ministry of Industry and Energy and in part by the Norwegian Ministry of Foreign Affairs.

References

Akerlof, G.A. (1991). "Richard T. Ely Lecture. Procrastination and Obedience." *American Economic Review* **81**(2):1-19.

Argyris, C. (1990). *Overcoming Organizational Defenses: Facilitation Organizational Learning*. Boston, MA: Allyn and Bacon.

Breivik, E., and Supphellen, M. (2003). "Elicitation of product attributes in an evaluation context: A comparison of three elicitation techniques." *Journal of Economic Psychology* **24**(1):77-98.

Brottemsmo, J.A., and Moxnes, E. (1991). "Automobile Choice: Determinants of Choice and Policies to Improve Fuel Efficiency." CMI-91-A30024. Bergen: Chr. Michelsen Institute.

Claxton, J.D., Fry, J.N., and Portis, B. (1974). "A Taxonomy of Prepurchase Information Gathering Patterns." *Journal of Consumer Research* **1**(December):35-42.

Dommermuth, W.P. (1965). "The shopping matrix and marketing strategy." *Journal of Marketing Research* **2**(May):128-32.

Fisher, A.C., and Rothkopf, M.H. (1989). "Market Failure and Energy Policy. A Rationale for Selective Conservation." *Energy Policy* (august):397-406.

Gately, D. (1980). "Individual Discount Rates and the Purchase and Utilization of Energy-Using Durables: Comment." *The Bell Journal of Economics* **11**:373-374.

Green, P., and Srinivasan, V. (1990). "Conjoint Analysis in Marketing: New Developments With Implications for Research and Practice." *Journal of Marketing* (October 1990):3-19.

Green, P.E., and Srinivasan, V. (1978). "Conjoint Analysis in Consumer Research: Issues and Outlook." *Journal of Consumer Research* **5**(Sept. -78):103-123.

Greene, D.L. (1998). "Why CAFE worked." *Energy Policy Energy Policy* **26**(8):595-613.

Gupta, P., and Ratchford, B.T. (1992). "Estimating the Efficiency of Consumer Choices of New Automobiles." *Journal of Economic Psychology* **13**:375-397.

Hatlebakk, M., and Moxnes, E. (1996). "Estimating the economic gain of energy standards on refrigerators." SNF-report 79/96. Bergen: SNF.

Hey, J.D. (1981). "Are optimal search rules reasonable? and vice versa? (And does it matter anyway?)." *Journal of Economic Behavior and Organization* **2**:47-70.

Kogut, C.A. (1990). "Consumer Search Behavior and Sunk Costs." *Journal of Economic Behavior and Organization* **14**(3):381-392.

Koomey, J.G., Mahler, S.A., Webber, C.A., and McMahon, J.E. (1999). "Projected regional impacts of appliance efficiency standards for the US residential sector." *Energy Energy* **24**(1):69-84.

Nadel, S. (2002). "Appliance and equipment efficiency standards." *Annual Review of Energy and the Environment Annu. Rev. Energ. Environ.* **27**:159-192.

Newman, J.W., and Staelin, R. (1972). "Prepurchase Information Seeking for New Cars and Major Household Appliances." *Journal of Marketing Research* **9**(august):249-257.

Ruderman, H., Levine, M.D., and McMahon, J.E. (1987). "The Behavior of the Market for Energy Efficiency in Residental Appliances Including Heating and Cooling Equipment." *The Energy Journal* **8**(1):101-124.

Sutherland, R.J. (1991). "Market Barriers to Energy-Efficiency Investments." *Energy Journal* **12**(3):15-34.

Sutherland, R.J. (1996). "The economics of energy conservation policy." *Energy Policy Energy Policy* **24**(4):361-370.

Turiel, I., McMahon, J., and Lebot, B. (1993). "Global Residental Appliance Standards." Pp. 183-191 in *The Energy Efficiency challenge for Europe. Proceedings of the 1993 ECEEE Summer Study. June 1-5*, edited by Ling, R. and Wilhite, H. Rungstedgård, Denmark.