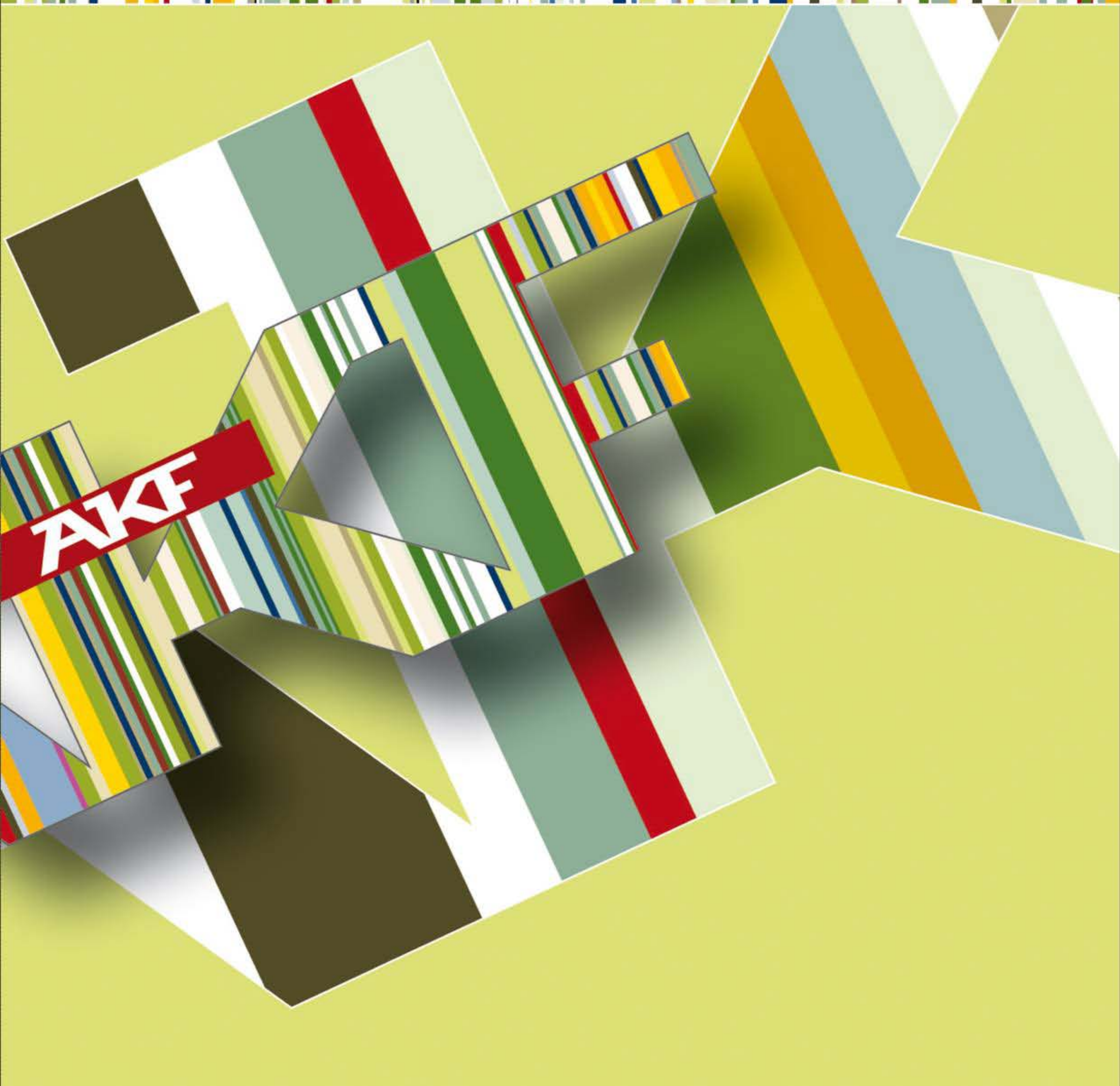
AKF Workingpaper

Vibeke Hansen Kjærbye

# Does Energy Labelling on Residential Housing Cause Energy Savings?



Does Energy Labelling on Residential Housing Cause Energy Savings?
Financial support from “Energistyrelsen” – Danish Energy Authority – is acknowledged.

AKF’s publications are available at www.akf.dk

AKF, Danish Institute of Governmental Research
Nyropsgade 37, DK-1602 Copenhagen V
Phone: +45 43 33 34 00
Fax: +45 43 33 34 01
E-mail: akf@akf.dk
Internet http://www.akf.dk



Publisher: AKF
ISBN: 978-87-7509-893-4
i:\forlaget\vha\energy labelling\energy_labelling.docx
December 2008(17)

**AKF, Danish Institute of Governmental Research**
Carries out and reports social science research of interest to the public sector and in particular to regions and local governments.

## Abstract

More than 80% of energy used in households is dedicated to space heating. Large potential energy savings have been identified in the existing housing stock. Energy labelling of single-family houses is seen as an important instrument to provide new house owners with information on efficient energy saving investments that can be made on the house. This paper evaluates the effects of the Danish Energy Labelling Scheme on energy consumption in existing single-family houses with propensity score matching using actual consumption of energy and register data describing the houses and households. We do not find significant energy savings due to the Danish Energy Labelling Scheme.



**Keywords**: Energy labelling, energy efficiency, propensity score matching

**JEL classification**: D12, Q41

# Contents

# 1 Introduction

Since the 1970s focus in the OECD countries has been on diminishing energy consumption. Nationally and internationally Denmark has made commitments to develop a more energy efficient consumption. Historically the reasons for reducing energy consumption have been to be less dependent on foreign fossil fuels and to lower the vulnerability to increasing energy prices. Today arguments for reducing energy consumption relate mainly to the climate change debate.

Denmark is one of few countries, which has experience with mandatory energy labelling schemes.[1] In Denmark, Energy Labelling Schemes have been used as a feasible mechanism to achieving energy savings in existing buildings since 1996. In 2003 EU introduced a directive on Energy Efficiency of Buildings which was strongly inspired by the Danish Energy Labelling Scheme for small buildings. The directive requires that all EU countries introduce energy labelling for buildings upon sale and rent.

In the US energy labelling is also used to achieve energy efficiency in buildings. A voluntary labelling scheme (RESNET[2] ) rated more than 165,000 new homes in 2006, representing approximately 10% of the US housing (see RESNET homepage 2008). The American RESNET standards have also been adopted in Canada and in the City of Shanghai, China.[3]

The purpose of this paper is to evaluate the effects of the Danish Energy Labelling Scheme on energy consumption in existing single-family houses, in order to determine whether *the Danish Energy Labelling Scheme for Small Buildings has caused significant energy savings*. Building related energy conservation has been a priority in Danish energy policy during three decades and is becoming increasingly central in both the EU and the US. The results of the research are expected to be useful in designing future policies in that respect.

This paper contributes to the scarce empirical literature by evaluating the effect of the Danish Energy Labelling Scheme on single-family houses with propensity score matching. We use actual consumption of energy (natural gas used for heating) and a very wide range of house and household characteristics obtained from administrative register data. This allowing us to assume Strong Ignorability, i.e. we control for all confounding variables determining energy labelling and energy consumption. In contrast to other estimation techniques matching is able to control for selection bias without imposing a particular parametric model for energy consumption.

As for general evaluations of energy efficiency in buildings, there is a comprehensive literature. A simple Google[4] search for "energy efficiency in buildings" gives more than 335,000 hits. And even though there is a growing interest in buildings' energy improvements, green building ratings and eco-labelling (see e.g. Building Energy Labelling Forum Agenda 2006) to our knowledge there are no micro-econometric evaluations of labelling schemes for buildings as tools to develop more energy efficiency in buildings worldwide.

---

1 Germany has had a labelling scheme since 1995, but has not collected data from the scheme centrally. Belgium has voluntary energy audits (see Thomsen et al. 2006).

2 Residential Energy Service Network is a national standards making body for building energy efficiency rating systems. Buildings are rated and a set of rated recommendations for cost-effective improvements that can be achieved by the rated building is also produced (see RESNET homepage 2008).

3 The World Bank estimates that by 2015, half of the world's new building constructions will take place in China, and more than half of China's urban residential and commercial stock will have been constructed after 2000 (see RESNET homepage 2008).

4 Search at scholar.google.com

The Danish scheme for energy labelling of buildings is based on a 1996 act, e.g. the "*Act to promote energy and water savings in buildings*", which among other initiatives sets out the rules for the Energy Labelling of Small Buildings (<1500 m²) and the Energy Management Scheme for Large Buildings (>1500 m²). Since the implementation of these schemes, relevant data in relation to the energy schemes have been collected for a database run and controlled by the secretariat for energy labelling under the Danish Energy Authority (Energistyrelsen).

Energy labelling is carried out for 45,000-50,000 single-family houses every year. More than 300,000 small buildings have been assessed over the first 6½ years, i.e. nearly 20% of all single-family houses in Denmark. In 2002 the total annual costs of the energy labelling scheme amounted to more than 90 million DKK (or about 12.8 million EUR) (see Laustsen and Lorenzen 2003).[5]

However, in spite of the high cost of energy labelling, there is very scarce evaluation evidence on the efficacy of labels to induce the recommended energy related investments and therefore enhance energy savings in the labelled houses. In 2001 an evaluation of the energy labelling scheme was carried out by Madsen, Ramlau and Pedersen (2001). This study found that energy related investment and energy savings in labelled and non-labelled houses were very similar. Two important limitations of this evaluation study were the small sample employed and the use of telephone interviews to obtain information.

International environmental labelling has been drawing much attention of policy makers and experts, but few studies on its ex post effectiveness have been undertaken. OECD highlights the lack of relevant evidence that labelling schemes significantly improve the energy performance of buildings (see OECD 2003).

Denmark has unique registers for both persons and buildings that can be combined with energy consumption data provided by the utilities. We observe a very wide range of house characteristics for both labelled and non-labelled houses, and this makes it possible to assume Strong Ignorability and use Counterfactual Framework to estimate causal effect. We connect energy labelling data with household energy consumption data and individual/household specific data and use matching analyses (see Rubin 1974; Rosenbaum and Rubin 1983) to evaluate the effect of the Danish energy scheme.

We estimate the average treatment effect on the treated of energy labelling for several years after the house is labelled. We do so separately for the three main categories of energy labelled houses A-, B- and C-labelled houses, since A-, B- and C houses imply different amounts of investments and in the hypothetic case that labelling induces savings we would expect different effects of the energy label according to the type of house.

The rest of this paper is organised as follows. Section 2 describes the Danish Energy Labelling Scheme, an earlier evaluation of the scheme and relevant literature. Section 3 provides a description of the data used in the analysis, highlighting the important differences between labelled and non-labelled houses and among the different categories of labelled houses. Section 4 describes the evaluation approach chosen in this paper based on the treatment effects model of Rubin (1974), and the particular estimator employed; a kernel propensity score matching estimator. Section 5 discusses the main empirical findings. The paper ends with a number of concluding remarks in section 6.

5 In 2008 the total annual cost amounted to 300-500 million DKK (or 42-71 million EUR) (see Ditlefsen 2008).

## 2 The Danish Energy Labelling Scheme

In Denmark, several instruments are used to reduce energy consumption for space heating in residential houses and to make the house stock more energy efficient. One of these instruments is energy labelling. Energy labelling has been mandatory in Denmark since 1997 for all existing buildings used for residential, public, trade or private services purposes, with two different schemes for large and small buildings. The background for this energy labelling scheme was to simplify and replace the former legislations, which had been evaluated as being confused and having questionable effect. This paper concentrates on the effect of the Energy Labelling Scheme for Small Buildings on single-family houses.

Energy labelling for buildings smaller than 1,500 m² includes energy rating, energy planning including documentation, detailed registration of building and installations and calculated consumption. Different from big houses, small houses are only labelled when they are put on the market for sale.[6] The main target for this scheme is single-family houses.

The main objective of the energy labelling scheme is to initiate energy and water savings in the housing stock by addressing consumers with factual information showing new owners or potential buyers that energy costs will account for a large part of the future costs. The scheme implies that all sold houses should be subject to an energy labelling so that buyers of a house receive information on the state of the energy installations, thermal insulations and water consuming equipment of the particular house and are further being presented with suggestions on possibilities of saving energy and water. All costs of the energy labelling are paid by the house seller, who has to pay the consulting engineer or architect for the energy labelling, including the energy audit and the necessary calculations. The typical price of a single-family house energy label is 2,000-3,500 DKK (or about 300-500 EUR).[7] Energy labelling is carried out by an approved energy consultant, who must have at least five years of documented relevant experience in building technology and energy consultancy.

The energy label consists of an energy label category and energy plan including documentation of the present state of the building, the heating system, the use of energy under the present owner, information on the expected use of the building and typical conditions such as price of energy, heating, size of household etc. Further information is given for every part of the building, heating system, automatics and for ventilation comprising information on type and the present insulation or energy efficiency etc. Finally, remarks on special situations or problems and comments are included. The label contains a standardised energy rating of the building containing information about the state of respectively heat, electricity and water installations and subsequently $CO_2$ emission impact. Findings are compared to other buildings with a similar use and based on number of residents, and subsequently the calculated consumption is placed on a scale from A1 to C5 (A: low, B: middle and C: high).

A1 will typically be given to a new built low-energy house or houses with similar energy conditions. A single-family house from the 1970s will typically attain a grade in the lowest A's or in the top of the B scale. A well energy-wise maintained house from the 1950s will typically get a grade from the B scale, while a half-timbered house with thatched roof built before 1900

---

6 Energy labelling for large buildings, e.g. buildings bigger than >1,500 m², is done on an annual basis and includes: energy rating, energy management, registration of consumption and energy planning including recommendations for improvements supported by calculation of the expected investment cost, annual heat and cost savings as well as life time of investment and pay-back period for potential investment. Data on net heating demand for room heating, electricity and water consumption are also presented.

7 The price of a single family house energy label was 2,000-3,500 DKK in 2002.

typically gets a grade from the C scale. Houses in the different labelling categories are quite different, not only in terms of proposed profitable saving possibilities, but also in terms of house characteristics, and this paper will evaluate separately the average effects of energy label for the three main labelling categories.

The energy plan includes a proposal for profitable saving possibilities for all types of energy and water consumption facilities of the building. Furthermore, the energy plan includes estimates of necessary investments and annual savings of the proposal. The plan also reports the estimated technical lifetime of the proposal and provides the necessary details for calculating how profitable the individual proposals are under a given financing. A proposal is defined as profitable if: ((annual savings in DKK * estimated technical lifetime)/estimated necessary investment in DKK) >1.33. This part of the label gives the potential buyer information about the house that would have been hidden for him if the energy labelling scheme had not existed. The information in itself does not give reason for any reduction in energy consumption.

It is important to stress that only if the new house owner carries out some or all of the proposed profitable energy saving improvements, the labelling scheme might induce future energy savings. Hence a rational owner will carry out a particular energy saving recommendation if the present value of the energy consumption plus the cost of the improvement are lower than the present value of the energy consumption without the energy investment.

As expenses for heating normally make up the largest part of the expenses for energy consumption in a household[8], the energy labelling scheme has special focus on improvements that can reduce these expenses. Madsen, Ramlau and Pedersen (2001) examined the proposed profitable saving possibilities registered in the energy labelling database and concluded that 95% of the recommended investments in energy saving improvements are related to reducing energy consumption for heating.

The energy labelling scheme is mandatory to existing houses in relation to a sale, but there are no significant consequences when the labelling is avoided. In the studied period, only 50-60% of the potential buildings were labelled when they were sold with important geographic differences in terms of the coverage of the label. Buildings in Greater Copenhagen and Funen are registered with coverage on more than 85%, while buildings in Northern Jutland are underrepresented with coverage on only 15-25% (see Madsen, Ramlau and Pedersen 2001).

According to the Danish Energy Authority 45,000-50,000 labels are issued each year. In total, more than 300,000 buildings corresponding to nearly 20% of all single-family houses in Denmark have got an energy label in the first 6½ years of the scheme. In 2001, energy savings for more than 130 million EUR were identified. The result of implementing all the possible savings would reduce the annual consumer energy cost by almost 20 million EUR. On average, the single-family houses could lower their energy costs by about 20%. More than 45% of the owners of labelled houses reveal that they have implemented energy savings in the first year they own a new house, but this is not alone due to the energy labelling of the house (see Laustsen and Lorenzen 2003).

In 2000 an evaluation of the Danish Energy Labelling Scheme for small buildings was carried out by the consultancy firm, COWI, by telephone interviews of 600 recent house buyers on their investments in energy savings – 300 energy labelled households and 300 non-labelled households were interviewed. This study found that only about 43% of the interviewed house owners knew or had heard about the scheme. Further, only half of the owners of labelled houses were familiar with the existence of the labelling of their house/flat. In-

[8] Other expenses are on electricity and water.

vestment levels and energy savings in labelled and non-labelled houses were compared and it was found that the differences were very small and almost statistically insignificant.

Yet, there were indications that labelling had made a difference in what improvements were made. Labelled houses had been subject to more technical demanding improvements with a larger saving potential, whereas non-labelled houses had made more aesthetic improvements like change of windows. The evaluation concluded that a larger sample would be necessary to show any possible significant differences in investment levels and energy savings (see Madsen, Ramlau and Pedersen 2001). The present paper will use a rather different evaluation method, a larger sample and be based on actual energy consumption and register data rather than respondents' answers to interviews. The partial coverage of the label allows us to construct a control group of non-labelled houses, and use matching estimation to evaluate the label effect.

Other papers have looked at the Danish labelling scheme, e.g. Jensen (2004) who identified a number of barriers for realising energy savings in buildings and Gram-Hanssen and Jensen (2006) who interviewed 10 families who had bought houses within the last three years, and found that the respondents remembered the label, they found that the labelling scheme was a good idea, but they had not really used the information from the label.

# 3 Data

For this study a comprehensive database for the period 1999-2002 has been constructed merging data from the energy labelling database, data from administrative registers containing socioeconomic data on house owners, e.g. age, education, income, family composition, data from public administrative registers (BBR) describing each house (both labelled houses and control group houses without an energy label), e.g. size, age, number of rooms, number of storeys and data on exact energy consumption (natural gas for heating) for two geographically different areas in Denmark provided by two natural gas companies (HNG – urban area close to the capital and MidtNord – a more rural area).

The initial main dataset had information on 74,348 single family houses (63,755 houses from the HNG area and 10,593 houses from the MidtNord area). A comprehensive work has taken place in order to impose a frame and exclude missing observations in the dataset.

First, house owners living together with their parents are excluded. Compared to a normal family composition such a household will consist of more than two adults, which can both affect energy consumption and household income, which again can affect energy consumption. Also very old houses (built before 1900) are excluded. This, because old houses can have a significant different insulation standard compared to more recently built houses. Significant bigger houses (>350 m²) are excluded under the assumption that energy use in very big houses is not comparable to energy use in average size houses. This initial trimming reduces the dataset to 47,099 houses.

In the reduced dataset 60% of the houses from the HNG dataset have no registered natural gas consumption. These are eliminated from the dataset since the estimations are made in relation to the use of natural gas. The dataset now has 22,872 observations. Finally, only houses traded in 1999-2002 are kept in the dataset. The final total dataset includes 3,956 single-family houses.

After this trimming all houses in the final dataset are between 50-350 m² and they are built after 1900. The houses are single-family houses used for residence only and occupied by the owner of the house. All houses use natural gas for heating and are located either in the HNG area or MidtNord area. All house owners are at least 18 years old and are not living together with their parents.

Of the 3,956 single-family houses in the final database 2,059 houses are energy labelled and 1,897 are non-labelled houses. For each house we have between 1 and 4 observations, as a house is observed once a year, and some houses enter the dataset later than 1999, and some leave before 2002. A house is defined as labelled, if the energy label is registered within one year after the house is sold. The control group houses are non-labelled houses sold between 1999 and 2002.

The dataset is non-random in at least two respects. First, even though labelling is compulsory a household faces no significant penalty by not entering the scheme. The buildings having entered it may therefore be different from the rest of the building population. Second, the availability of energy consumption data, in this case natural gas for heating, is first limited to major city areas, where natural gas heating is provided, and secondly to areas, where the natural gas companies are willing to provide household natural gas consumption data for research purposes.

Earlier evaluations of the energy labelling scheme have pointed out that in the Greater Copenhagen and Funen area more than 85% of all sold smaller buildings are energy labelled, while energy labelled buildings in Northern Jutland are underrepresented with a coverage of only 15-25% of all sold smaller buildings (Laustsen and Lorenzen 2003). In this paper the dataset are constructed to handle differences in geographical distributions of the coverage in energy labelling by including natural gas consumption data from two geographically different natural gas companies. Table 1 shows an increase in the coverage from 0.42 to 0.62 during the observed period, meaning that an increasing percentage of the traded houses gets labelled.

**Table 1 Development in coverage of the labelling scheme**

| | 1999 | 2000 | 2001 | 2002 | Total |
|---|---|---|---|---|---|
| Annual number of houses getting labelled | 425 | 533 | 585 | 516 | 2059 |
| Total number of house purchases per year in the dataset | 1006 | 1043 | 1079 | 828 | 3956 |
| % coverage (labelled houses/bought houses) | 0.42 | 0.51 | 0.54 | 0.62 | 0.52 |

Section 2 described how the energy labelling scheme gives houses grades in the range from A1-C5. Table 2 shows the distribution of houses in the three main categories A, B and C, and the control group in relation to number of years since the house was bought. As seen, the main percentage of the labelled houses are B-houses (59%), 21% are A-houses and 20% C-houses. Most houses are observed in the first and second year after the house is bought.

**Table 2 Distribution of houses in the control group and three main labelling categories in relation to number of years since the house was bought**

| | Year of house purchase (Ysb=0) | Ysb=1 | Ysb=2 | Ysb=3 | Ysb=4 | Total |
|---|---|---|---|---|---|---|
| A | 53 | 327 | 273 | 164 | 75 | 892 |
| B | 100 | 983 | 748 | 444 | 208 | 2483 |
| C | 28 | 325 | 274 | 164 | 61 | 852 |
| Controls | 545 | 1045 | 1144 | 712 | 366 | 3812 |
| Missing information | | | | | | 10 |
| # observations | 726 | 2680 | 2439 | 1484 | 710 | 8049 |

Our covariate set includes a range of house and household characteristics that might be correlated with labelling propensity and influence consumption of energy for heating. As house characteristics we consider house size and house size related controls like number of toilets, number of bathrooms, number of floors, house structural characteristics like type of roof, type of outside wall, and type of heating installation, and finally year of construction.[9] As household observable characteristics we include number of children at different age intervals, age and education of the main person, and household disposable income.

A variable describing the outdoor temperature is also necessary, when we estimate energy consumption for heating. Here the variable "graddag" (degree day) describes how cold the weather has been each year. An increasing number of degree days means a year with colder mean temperature. Other unexplained annual variations affecting energy consumption are described by dummies for each year (D1999, D2000, D2001). If there should be any

9 Petersen and Gram-Hanssen (2005) point out house size and year of construction looked among the most relevant factors determining energy and water consumption in the Danish households.

structural changes in the house sales market they are caught by dummies describing the year of house purchase. To take account of geographical differences between the two areas we include a dummy variable (Mn=1 if the house is situated in the MidtNord area, and Mn=0 if the house is situated in the HNG area).

Table 3 reports descriptive statistics for the covariate set. Due to the fact that we do not observe all houses four years after the house is bought, this table includes the covariate distribution for the biggest of the four sub-samples used in the application, which is the sample of houses observed one year after the house is bought. It can be seen that a typical house is a brick house with slate roof. The house is linked to a central heating system based on natural gas. It has two toilets and one bathroom, it is a one storey house and there are four members of the household. The house was built 1960-1969.

Table 3 also highlights the differences and similarities between both the three labelling categories of labelled houses, and between labelled and control group houses. For example, we can see that A-labelled houses are bigger than control group houses, B-labelled houses are about the same size, and C-labelled houses are smaller. We also see that A- and B-labelled houses use less energy/$m^2$ for heating than control group houses, while C-labelled houses use more. As for construction materials used on the houses we see both differences and similarities between the groups. The majority of both A-, B-, C- and control houses fits the general description (brick house with slate roof and natural gas based central heating). Most C-houses have only one toilet, whereas A-, B- and control houses mostly have two toilets. The majority of houses in all sub-samples has one bathroom, but when comparing across the groups, a larger proportion of A-houses has two bathrooms than in the control group, the proportion of B-houses with two bathrooms is like the control group, and a smaller part of the C-houses has two bathrooms.

**Table 3 Descriptive Statistics for the first year after house purchase (Ysb=1) Presented numbers are means for the covariates**

| | Description of variable | A-label | B-Label | C-label | All labelled houses | Controls |
|---|---|---|---|---|---|---|
| House characteristics | | | | | | |
| Size of house | $m^2$ | 150.91 | 139.02 | 118.71 | 137.40 | 142.56 |
| Energy consumption | Kwh/ $m^2$ | 11.14 | 14.52 | 17.40 | 14.40 | 15.21 |
| Construction year | | 1970 | 1959 | 1945 | 1958 | 1956 |
| Roof | Slate roof | 0.44 | 0.53 | 0.42 | 0.49 | 0.47 |
| | Cement | 0.25 | 0.08 | 0.02 | 0.10 | 0.11 |
| | Tile | 0.21 | 0.25 | 0.42 | 0.28 | 0.29 |
| | Other roof material[1] | 0.10 | 0.14 | 0.14 | 0.13 | 0.13 |
| Outside wall | Brick | 0.91 | 0.87 | 0.74 | 0.85 | 0.87 |
| | Concrete | 0.04 | 0.10 | 0.17 | 0.10 | 0.09 |
| | Other outside wall material[2] | 0.05 | 0.04 | 0.09 | 0.05 | 0.04 |
| Heat installation | Central heating | 0.97 | 0.98 | 0.97 | 0.98 | 0.95 |
| | Other heating[3] | 0.03 | 0.02 | 0.03 | 0.02 | 0.05 |
| Supplementary heating | Wood burner | 0.21 | 0.23 | 0.14 | 0.21 | 0.19 |
| | Solar panel | 0.01 | 0.01 | 0 | 0.01 | 0.06 |
| | Open fireplace | 0.03 | 0.04 | 0.03 | 0.04 | 0.04 |
| | Other suppl. heating[4] | 0.03 | 0.02 | 0 | 0.02 | 0.02 |
| | No suppl. heating | 0.72 | 0.70 | 0.83 | 0.73 | 0.75 |
| # Toilets | 1 toilet | 0.23 | 0.39 | 0.61 | 0.40 | 0.39 |
| | 2 toilets | 0.69 | 0.56 | 0.34 | 0.54 | 0.54 |
| | 3-4 toilets | 0.08 | 0.06 | 0.05 | 0.06 | 0.07 |
| # Bathrooms | 1 bathroom | 0.53 | 0.70 | 0.82 | 0.69 | 0.69 |
| | 2 bathrooms | 0.46 | 0.29 | 0.14 | 0.29 | 0.29 |
| | 3-4 bathrooms | 0.02 | 0.01 | 0.01 | 0.01 | 0.02 |
| # Floors | 1 floor | 0.98 | 0.98 | 0.98 | 0.98 | 0.97 |
| | 2-3 floors | 0.02 | 0.02 | 0.02 | 0.02 | 0.03 |
| Vintage class | Built 1900 – 1949 | 0.11 | 0.22 | 0.48 | 0.25 | 0.30 |
| | 1950- 1959 | 0.05 | 0.14 | 0.29 | 0.15 | 0.14 |
| | 1960-1969 | 0.19 | 0.37 | 0.20 | 0.30 | 0.28 |
| | 1970-1976 | 0.30 | 0.24 | 0.02 | 0.21 | 0.18 |
| | 1977-1981 | 0.15 | 0.03 | 0 | 0.05 | 0.05 |
| | 1982-1984 | 0.05 | 0.003 | 0 | 0.01 | 0.004 |
| | 1985-1997 | 0.15 | 0.01 | 0 | 0.04 | 0.05 |
| | 1998-1999 | 0 | 0 | 0 | 0 | 0 |

| | Description of variable | A-label | B-Label | C-label | All labelled houses | Controls |
|---|---|---|---|---|---|---|
| | 2000 + | 0 | 0 | 0 | 0 | 0 |
| Household characteristics | | | | | | |
| Total household income | | 755,970 | 724,099 | 656,941 | 718,449 | 590,754 |
| Age of house owner at year of house purchase | | 37.81 | 36.62 | 35.31 | 36.59 | 37.00 |
| Number of members in household | | 3.35 | 3.34 | 3.19 | 3.31 | 3.29 |
| Number of children | Age 0-6 | 0.74 | 0.73 | 0.72 | 0.73 | 0.47 |
| | Age 7-14 | 0.33 | 0.35 | 0.25 | 0.33 | 0.31 |
| | Age 15-17 | 0.08 | 0.09 | 0.07 | 0.08 | 0.05 |
| Highest level of finished education for main person | Niv1 | 0.11 | 0.12 | 0.10 | 0.11 | 0.17 |
| | Niv2 | 0.45 | 0.45 | 0.42 | 0.44 | 0.43 |
| | Niv3 | 0.23 | 0.24 | 0.34 | 0.26 | 0.23 |
| | Niv4 | 0.20 | 0.17 | 0.12 | 0.16 | 0.12 |
| | Niv0 | 0.02 | 0.02 | 0.02 | 0.02 | 0.04 |

1 E.g. roofing felt, corrugated iron, thatched roof, fibre cement, PVC.

2 E.g. eternit, half-timbered, wood, concrete, fibre cement.

3 E.g. district heating, stove, heat pump, electricity inst., gas radiator.

4 E.g. heat pump, stove for liquid fuels, electric stove.

The construction year is quite different for A- and C-houses. The mean construction year for A-houses is 1970. The mean construction year for B-houses is 1959 and the mean construction year for C-houses is 1945. The mean construction year for control houses is 1956. As there has been a tightening of the building regulations over time, houses built before 1979 did not have to fulfil the same demands for e.g. insulation as houses built later. We would therefore expect more recently built houses to use less energy for heating than old houses, everything else being equal.

The socioeconomic characteristics of households living in the different types of labelled houses also differ. There are no significant age differences regarding the owner among the four types of houses, but households in labelled houses have more children between the age if 0-6, than households in control houses. The highest number of children in this category is seen in A-labelled houses, next are B-labelled houses and C-labelled houses. Also education varies between labelled and control households; households living in labelled houses generally have a higher education than households living in control group houses. This difference in education is reflected in household income, where we find the highest average household income in A-houses followed by B-houses, C-houses and the lowest average income corresponds to control houses.

The description of the dataset in this section points at relevant differences between labelled houses and non-labelled houses, but also among the three categories of labelled houses. Differences between house characteristics call for matching because it is important to balance properly control houses to the characteristics of the labelled houses. The differences in terms of covariate distribution justify using different propensity scores for each type of labelled houses; A, B and C.

This paper hypothesises that for energy labelling to show some efficiency, the average energy saving would be positive and at least not decreasing over time. This reflecting that recommended investments can be adopted immediately or gradually. Because A-labelled houses are the most energy efficient houses we further hypothesise the average energy saving of a C-house not to be smaller than the average energy saving of a B-house or an A-house, and that the average energy savings of a B-house not to be smaller than the average energy savings of an A-house.

# 4 The Evaluation Approach

The objective of this paper is the evaluation of the Danish Energy Labelling Scheme on household energy consumption. The estimation of the labelling scheme impact should ideally be based on experimental data. However, due to the unavailability of such data, the paper uses non-experimental comparison groups of labelled and non-labelled houses and estimates the effect of labelling on labelled houses with propensity score matching (see Rosenbaum and Rubin 1983 or Dehejia and Wahba 1999).

The treatment effect model (see Rubin 1974) has been extensively applied by labour economists to assess the impact of active labour-market policies (see for example Heckman and Robb 1985; Heckman, Ichimura and Todd 1997). This paper will use this framework to estimate the average treatment effect on the treated (ATT) parameter which in our case is the average labelling effect in terms of natural gas consumption on labelled houses.

The treatment effect model assumes that each house is potentially exposable to labelling, such that we can think of two potential states for each house. In state 0 (denoted $D$=0) a house is not labelled, while in state 1 a house is labelled. Each state has an associated outcome referred to as the potential outcome and denoted $Y^0$, $Y^1$. The parameter of interest, the average labelling effect for a labelled house is the expected difference of potential outcomes for this type of house, denoted ATT=$E(Y^1 - Y^0|D=1)$.

The fundamental problem of the treatment effects framework is that $Y^1$ and $Y^0$ are never jointly observed for a particular house, and therefore we are forced to rely on comparisons between labelled and non-labelled houses. Concretely, the assumption of ignorability of treatment, also called selection on observables which in its weakest version requires mean independence of $Y^0$ and D given observable characteristics $X$ (see Rosenbaum and Rubin 1983): [10]

$$E(Y^0|X,D=0) = E(Y^0|X,D=1) = E(Y|X)$$

the parameter of interest, ATT is identified.

Under this assumption, ATT is obtained from the observable conditional means $E(Y|D=1,X)$ and $E(Y|D=0,X)$, since ATT=$E(ATT(X)|D=1)$ with ATT($X$)= $E[Y^1 - Y^0|X,D=1] = E[Y^1|X,D=1] - E[Y^0|X,D=1] = E[Y|X,D=1] - E[Y|X,D=0]$.

The assumption of ignorability-of-treatment implies that we observe all variables, X, that jointly determine selection into treatment and potential outcome without treatment. In our application, this set of confounders includes a wide range of house and household specific characteristics.[11]

Rosenbaum and Rubin (1983) show that when matching on $X$ is valid, matching on the probability of being treated given confounders $X$, called the propensity score $p(X)=P(D=1|X)$, is also valid:

---

[10] For estimating the ATT parameter, matching methods allow selection into treatment to be based on possibly unobserved components of the anticipated programme impact, but only insofar that the programme participation decisions are based on the unobservable determinants of $Y^1$ and not those of $Y^0$. The matching method also requires that the distribution of the matching variables is not affected by whether the treatment is received. Variables that are likely to be affected by the treatment are used in the set of matching variables.

[11] See section 3

$$E(Y^0|P(X),D=0)= E(Y^0|P(X),D=1)=E(Y|P(X))$$

Much of the matching literature focuses on propensity score matching methods (PSM), because adjustment for the propensity score suffices for removing all biases associated with differences in the covariates between treated and controls. When using PSM of Rosenbaum and Rubin (1983), the matching procedure is broken down into two stages. The first stage estimates the propensity score $P(D=1|X)$ using a binary discrete choice model. The second stage matches houses on the basis of their predicted probabilities of being labelled.

Many different PSM methods can be obtained by using different methods at the first and second stage. This paper estimates the propensity score with a probit model and uses kernel matching method at the second stage. The idea of this estimator is to match the energy consumption of each labelled house with a weighted average over the set of all houses in the control group:

$$A\tilde{T}T = N_1^{-1}\left(\sum_{i=1}^{N_1}(Y_i D_i - \tilde{Y}_{0i})\right)$$

$$\tilde{Y}_{0i} = \sum_{j=1}^{N_0} w(i,j)(1-D_j)Y_j$$

where the weights associated to control j regarding treated i, $w(i,j)$, have an Epanechnikov form and their magnitude depends on the distance between predicted probability of control $p_j$ and treated house $p_i$:

$$w(i,j) = \frac{G(\frac{p_j - p_i}{b})}{\sum_{k=1}^{N_0} G(\frac{p_k - p_i}{b})},$$

where b is the bandwidth (see Heckman, Ichimura and Todd 1997).

# 5 Results

This section discusses the results from estimating the average effect on future natural gas consumption of the Danish Energy Labelling Scheme with PSM. The PSM analysis will be performed separately for three sub-samples. Section 3 revealed big differences in terms of covariate distribution not only between labelled and control houses, but also between different labelling categories. In addition, if the labelling scheme enhances energy savings, we would expect important differences among A-, B- and C-labelled houses. A-labelled houses by definition are the best rated houses in relation to energy efficiency, and C-labelled houses are the worst, we would expect C-labelled houses to get the largest number of recommended improvements, and thereby have the greatest potential to improve the energy efficiency for the house. In other words, we would expect there to be largest effect of the energy labelling scheme for C-labelled houses. In order not to obscure potential different effects, the analysis is performed separately for three samples, where the treatment group is composed by A-, B- or C-labelled houses and the control group is composed by non-labelled houses.

Furthermore, there might be differences in the effect of energy labelling according to how many years the house owner has had to carry out the recommended improvements from the energy label, since some of the recommended investments might be postponed to future years[12]. The simple fact that a house is labelled with an energy label does not have any effect on a household's energy use. Only when the house owner implements some or all of the recommended improvements suggested in the energy label (Energy Plan), the labelling scheme will have an effect when comparing labelled houses to control group houses.

Three general hypothetical scenarios can be thought of in relation to the house owner's implementation of these recommendations. First, the house owner can totally ignore the recommendations or at least deliberately decide not to implement any of them. This will result in no significant difference in energy use compared to the control group. Second, the house owner decides to implement all recommendations the first year of owning the house and therefore uses less energy from year 1. This will result in a constant energy saving in all years. Third, the house owner decides to implement the recommendations gradually, resulting in an increasing energy saving across future years.

An example of change in behaviour that could lead to insignificant ATT even though the house owner has implemented some or all recommendations could be: the house owner implements some recommendations and can obtain the same indoor temperature as before by less consumption of energy and he thereby saves money in the form of a smaller energy bill. The money he saves is used to increase consumption of energy to raise the indoor temperature, because it gives him a higher utility. Hereby the energy label and its recommendations actually lead to an increase instead of the intended decrease in energy used for heating. But at the same time the energy label also leads to an increase in utility for the house owner.

Because of the expected effect differences in A-, B- and C-labelled houses and expected differences in effect over time since the house was bought, the initial dataset is broken down into 12 sub-samples (A-, B- and C-labelled houses in 1-4 years after house purchase) and separate propensity scores, matches and ATT are estimated for each sub-sample.

The propensity scores are estimated with a probit model, this leading to the predicted probability of a house being labelled given observable heterogeneity. We impose common support by excluding those labelled houses whose predicted probability is outside the range of the predicted probabilities for the control group.

[12] The recommended improvements are presented in the part of the energy label called Energy Plan.

First, we present evidence on the matching estimators' ability to balance the covariates for each of the sub-samples. The balancing test results are presented in tables 4, 5, 6 and 7. Table 4 presents the results for the sub-samples of A-, B- and C-labelled houses and their natural gas consumption one year after the house purchase. Table 5 presents results for the sub-samples two years after the house purchase and so forth. The *SDIFF before* match estimates in columns 2, 5 and 8 are the standardised difference between the sample means of the labelled houses and the control group.

**Table 4 Covariates balance analysis between labelled and non-labelled houses based on Epanechnikov kernel matching. Dependent variable = log (consumption of natural gas / size of house). Year 1 after house purchase**

| | Matched A-labelled houses | | | Matched B-labelled houses | | | Matched C-labelled houses | | |
|---|---|---|---|---|---|---|---|---|---|
| # labelled houses | 317 | | | 973 | | | 324 | | |
| # control houses | 1045 | | | 1045 | | | 1045 | | |
| Variable name | SDIFF before match | SDIFF after match | p>\|t\| | SDIFF before match | SDIFF after match | p>\|t\| | SDIFF before match | SDIFF after match | p>\|t\| |
| Mn | -39.3 | -0.1 | 0.980 | -36.3 | 0.7 | 0.783 | -30.6 | -1.1 | 0.846 |
| Opfort_1949 | -46.6 | 0.0 | 0.994 | -18.9 | -1.1 | 0.791 | 37.5 | -0.1 | 0.989 |
| Opfort_1959 | -28.9 | -0.7 | 0.904 | -0.3 | 1.4 | 0.758 | 38.4 | -0.7 | 0.939 |
| Opfort_1976 | 30.4 | -2.3 | 0.793 | 15.1 | -2.3 | 0.634 | -56.0 | -2.2 | 0.579 |
| Opfort_1981 | 34.7 | 6.4 | 0.492 | -13.3 | -0.9 | 0.822 | - | - | - |
| Opfort_1984 | 12.1 | -6.2 | 0.572 | -1.3 | 0.8 | 0.855 | - | - | - |
| Opfort_1997 | 35.1 | 1.9 | 0.840 | -23.3 | 0.6 | 0.814 | - | - | - |
| Lntotindk | 49.4 | 1.5 | 0.763 | 43.6 | -0.6 | 0.815 | 25.3 | -0.5 | 0.938 |
| Lnboligarl | 31.9 | -4.0 | 0.577 | -4.5 | -1.6 | 0.722 | -69.2 | -1.2 | 0.875 |
| A_yob | 7.9 | -1.3 | 0.860 | -2.9 | -3.2 | 0.437 | -15.1 | 0.8 | 0.907 |
| Antper_a | 3.1 | 2.3 | 0.756 | 3.0 | -1.5 | 0.717 | -8.4 | -2.2 | 0.748 |
| Antper2 | -2.9 | 1.1 | 0.865 | -1.8 | -1.1 | 0.783 | -12.5 | -1.7 | 0.776 |
| Anc_06 | 29.6 | 4.3 | 0.615 | 32.2 | -1.3 | 0.793 | 31.2 | 0.6 | 0.943 |
| Anc_714 | 5.1 | -1.7 | 0.829 | 6.1 | -0.9 | 0.843 | -9.9 | -4.1 | 0.587 |
| Anc_1517_a | 11.5 | 1.7 | 0.845 | 12.7 | -0.5 | 0.928 | 8.7 | -0.1 | 0.986 |
| Lbeton | -21.1 | 2.2 | 0.714 | -0.1 | 0.3 | 0.954 | 21.8 | -1.7 | 0.844 |
| Andet_y | 3.5 | 3.1 | 0.700 | -0.1 | 0.4 | 0.927 | 16.9 | -0.4 | 0.967 |
| Cement | 30.7 | 1.3 | 0.882 | -11.5 | -0.4 | 0.920 | -35.4 | 0.0 | 0.999 |
| Tegl | -17.2 | 1.2 | 0.874 | -9.1 | -1.7 | 0.701 | 26.9 | -2.7 | 0.740 |
| Andet_t | -10.0 | -1.6 | 0.834 | 3.1 | 1.3 | 0.775 | 0.8 | 0.7 | 0.928 |
| Andet_v | -10.1 | -0.9 | 0.900 | -17.4 | 0.0 | 0.999 | -12.2 | -1.2 | 0.858 |
| Andet_o | -26.0 | -0.1 | 0.987 | -20.6 | -2.5 | 0.520 | -9.1 | 1.4 | 0.846 |

| | Matched A-labelled houses | | | Matched B-labelled houses | | | Matched C-labelled houses | | |
|---|---|---|---|---|---|---|---|---|---|
| # labelled houses | 317 | | | 973 | | | 324 | | |
| # control houses | 1045 | | | 1045 | | | 1045 | | |
| Variable name | SDIFF before match | SDIFF after match | p>\|t\| | SDIFF before match | SDIFF after match | p>\|t\| | SDIFF before match | SDIFF after match | p>\|t\| |
| Andet_s | 5.6 | -0.7 | 0.935 | -2.3 | -0.0 | 0.994 | - | - | - |
| Brandovn | 5.5 | -0.1 | 0.988 | 8.8 | -2.3 | 0.618 | -12.6 | -4.6 | 0.542 |
| Pejs | -3.7 | 1.8 | 0.805 | 2.0 | 1.5 | 0.741 | -5.9 | 1.8 | 0.802 |
| Solp | 7.2 | -0.1 | 0.995 | -2.3 | -0.6 | 0.893 | - | - | - |
| Toilet_1 | -34.9 | 2.1 | 0.769 | 0.4 | -1.2 | 0.791 | 45.4 | -3.8 | 0.630 |
| Toilet_3 | 3.4 | 0.7 | 0.931 | -2.5 | -0.9 | 0.841 | -8.6 | 0.9 | 0.904 |
| Bad_2 | 33.6 | -3.7 | 0.654 | -1.4 | 1.3 | 0.774 | -37.6 | -0.5 | 0.945 |
| Bad_3 | 0.4 | -0.4 | 0.963 | -2.5 | -2.4 | 0.590 | -5.5 | 3.9 | 0.516 |
| Etager_2 | -2.5 | 1.8 | 0.808 | -6.5 | 0.3 | 0.933 | -2.8 | -1.8 | 0.813 |
| Niv1 | -17.6 | -2.3 | 0.755 | -13.2 | 1.3 | 0.761 | -21.3 | -1.0 | 0.889 |
| Niv3 | 1.2 | -0.6 | 0.938 | 2.1 | -1.4 | 0.762 | 22.0 | 3.0 | 0.714 |
| Niv4 | 14.6 | 1.6 | 0.849 | 12.0 | -1.5 | 0.758 | -0.3 | -2.4 | 0.761 |
| Niv0 | -16.6 | -2.1 | 0.732 | -12.6 | -2.2 | 0.575 | -12.6 | -0.7 | 0.914 |
| Graddag | 21.5 | 0.9 | 0.908 | 22.1 | 3.0 | 0.490 | 30.0 | -1.6 | 0.914 |
| D2002 | 0.3 | -1.9 | 0.808 | 9.4 | 0.2 | 0.966 | 0.6 | -1.3 | 0.872 |
| Yob99_00 | -18.9 | 0.2 | 0.983 | - | - | - | 11.5 | -0.9 | 0.908 |

Note: p-value is the probability for the means of treated and controls to be equal. P-values are estimated by Hotellings t-test.

As seen in the previous section, there are important differences in terms of covariates between labelled and non-labelled houses reflected at quite big standardised differences before matching for many covariates. Note that for the same covariates like Mn, Lntotindk, Anc_06, or Graddag, the three types of labelled houses depart from control houses in terms of that covariate in the same direction, while for other covariates like Opfort_x, Lnboligarl, Lbeton, Cement, Toilet_1, B-houses are similar to the control houses, while A- and C-houses depart in different directions from the controls in terms of the covariate. This highlights the necessity of separate propensity score analysis for the three types of houses.

As seen in tables 4-7, the Epanechnikov kernel matching balances treatment and control samples quite well, with no covariate presenting SDIFF after matching bigger than 20%, the critical value for reasonable bias suggested by Rosenbaum and Rubin (1985), and almost all covariates with biases after matching smaller than 10%. When comparing across the three sub-samples, the *SDIFF after match* for the B-labelled houses are in general the smallest. The *SDIFF after match* for immediate years after the house is bought are generally smaller than for more distant years due to sub-samples getting smaller. Still none of the *SDIFF after match* gets close to 20%.

Summing up, the matching procedure performs very well for the different sub-samples. Even 4 years after the house purchase (see table 7), where we have a smaller number of observations, and the standardised differences are larger, none exceeds 20. The p-values of Hotellings t-test presented in column 4, 7 and 10 complement the picture offered by the Standardized Bias Difference.

**Table 5 Covariates balance analysis between labelled and non-labelled houses based on Epanechnikov kernel matching. Dependent variable = log (consumption of natural gas/size of house). Year 2 after house purchase**

| | Matched A-labelled houses | | | Matched B-labelled houses | | | Matched C-labelled houses | | |
|---|---|---|---|---|---|---|---|---|---|
| # labelled houses | 273 | | | 745 | | | 278 | | |
| # control houses | 1144 | | | 1144 | | | 1144 | | |
| Variable name | SDIFF before match | SDIFF after match | p>\|t\| | SDIFF before match | SDIFF after match | p>\|t\| | SDIFF before match | SDIFF after match | p>\|t\| |
| Mn | -42.3 | -5.1 | 0.513 | -70.6 | -2.1 | 0.596 | -67.0 | -4.6 | 0.496 |
| Opfort_1949 | -36.4 | -1.0 | 0.889 | -12.8 | 0.5 | 0.917 | 47.1 | 1.6 | 0.861 |
| Opfort_1959 | -32.6 | -0.1 | 0.987 | 0.2 | 2.2 | 0.672 | 33.9 | -3.5 | 0.721 |
| Opfort_1976 | 14.9 | 0.2 | 0.982 | 6.3 | -4.3 | 0.422 | -63.3 | -2.9 | 0.515 |
| Opfort_1981 | 32.7 | 5.8 | 0.553 | -8.8 | -3.4 | 0.486 | -29.5 | -3.2 | 0.542 |
| Opfort_1984 | 21.4 | 1.0 | 0.924 | -8.9 | -0.1 | 0.967 | - | - | - |
| Opfort_1997 | 29.5 | -1.1 | 0.909 | -27.0 | 1.1 | 0.722 | - | - | - |
| Opfort_1999 | 9.9 | -2.4 | 0.817 | - | - | - | - | - | - |
| Lntotindk | 39.6 | -0.5 | 0.951 | 35.7 | -2.1 | 0.631 | 20.7 | -2.7 | 0.733 |
| Lnboligarl | 38.4 | -4.8 | 0.557 | -5.7 | -1.7 | 0.743 | -59.3 | -1.0 | 0.907 |
| A_yob | -3.1 | 0.9 | 0.907 | -10.7 | -0.8 | 0.855 | -14.6 | -0.7 | 0.924 |
| Antper_a | 23.7 | -2.1 | 0.788 | 15.2 | -0.0 | 0.995 | 6.1 | -2.6 | 0.728 |
| Antper2 | 13.0 | -2.3 | 0.736 | 7.3 | -0.7 | 0.861 | -1.3 | -2.0 | 0.750 |
| Anc_06 | 30.1 | 0.1 | 0.989 | 27.9 | 0.7 | 0.896 | 23.6 | 0.6 | 0.947 |
| Anc_714 | 7.7 | -2.0 | 0.819 | 9.9 | -0.3 | 0.961 | -3.3 | -2.4 | 0.771 |
| Anc_1517_a | 5.9 | -6.6 | 0.478 | 5.1 | -3.4 | 0.537 | 3.0 | 0.8 | 0.928 |
| Lbeton | -28.3 | 0.4 | 0.937 | 4.8 | 1.8 | 0.735 | 24.4 | 2.8 | 0.774 |
| Andet_y | -1.6 | 2.9 | 0.713 | 0.2 | -0.0 | 0.993 | 11.8 | -4.4 | 0.653 |
| Cement | 30.2 | -0.4 | 0.969 | -11.0 | -2.9 | 0.552 | -43.1 | -1.7 | 0.732 |
| Tegl | -0.3 | -0.8 | 0.927 | -1.3 | 1.7 | 0.739 | 30.1 | 0.6 | 0.946 |
| Andet_t | -10.3 | 2.3 | 0.767 | 7.9 | 1.3 | 0.812 | 7.8 | 2.4 | 0.786 |
| Andet_v | 2.9 | 3.0 | 0.722 | -3.3 | 1.1 | 0.823 | -15.8 | 0.6 | 0.909 |
| Andet_o | -5.1 | 1.7 | 0.826 | -0.0 | -0.3 | 0.949 | 6.9 | 4.5 | 0.601 |
| Andet_s | 3.9 | 0.7 | 0.935 | -6.4 | -1.0 | 0.829 | - | | |

| | Matched A-labelled houses | | | Matched B-labelled houses | | | Matched C-labelled houses | | |
|---|---|---|---|---|---|---|---|---|---|
| # labelled houses | 273 | | | 745 | | | 278 | | |
| # control houses | 1144 | | | 1144 | | | 1144 | | |
| Brandovn | 0.0 | -3.0 | 0.725 | 11.2 | -1.8 | 0.749 | -6.5 | -5.9 | 0.489 |
| Pejs | 4.9 | 1.0 | 0.909 | 3.2 | 3.1 | 0.548 | -1.7 | 2.7 | 0.737 |
| Solp | 12.1 | -13.7 | 0.276 | 1.9 | -0.1 | 0.986 | - | | |
| Toilet_1 | -44.9 | 0.2 | 0.983 | -5.3 | -0.4 | 0.935 | 37.4 | -2.5 | 0.765 |
| Toilet_3 | 12.5 | -2.7 | 0.780 | -0.1 | -0.7 | 0.891 | -5.0 | 1.3 | 0.873 |
| Bad_2 | 38.7 | 1.6 | 0.862 | 2.2 | 2.2 | 0.667 | -32.5 | -0.8 | 0.915 |
| Bad_3 | 7.5 | -8.2 | 0.447 | -1.4 | -0.2 | 0.969 | 1.9 | 5.3 | 0.504 |
| Etager_2 | 5.0 | -0.3 | 0.978 | -0.7 | 0.4 | 0.936 | -0.3 | 0.1 | 0.986 |
| Niv1 | -14.7 | 0.1 | 0.986 | -18.6 | -1.1 | 0.823 | -20.8 | -3.9 | 0.614 |
| Niv3 | 5.5 | 0.4 | 0.967 | 2.7 | 1.1 | 0.835 | 23.6 | 3.2 | 0.717 |
| Niv4 | 14.9 | 1.4 | 0.877 | 14.9 | -0.1 | 0.984 | -0.7 | -3.7 | 0.669 |
| Niv0 | -5.4 | -3.8 | 0.647 | -7.0 | -1.2 | 0.803 | -3.0 | -7.2 | 0.426 |
| Graddag | 2.9 | -2.8 | 0.748 | 6.0 | 2.5 | 0.636 | -15.2 | -3.5 | 0.694 |
| D2001 | -23.0 | 2.8 | 0.733 | -18.4 | -2.4 | 0.632 | -30.9 | 0.7 | 0.933 |
| Yob99_00 | -18.4 | 4.4 | 0.615 | - | - | - | - | - | - |

**Table 6 Covariates balance analysis between labelled and non-labelled houses based on Epanechnikov kernel matching. Dependent variable = log (consumption of natural gas / size of house). Year 3 after house purchase**

| | Matched A-labelled houses | | | Matched B-labelled houses | | | Matched C-labelled houses | | |
|---|---|---|---|---|---|---|---|---|---|
| # labelled houses | 157 | | | 443 | | | 165 | | |
| # control houses | 712 | | | 712 | | | 712 | | |
| Variable name | SDIFF before match | SDIFF after match | p>\|t\| | SDIFF before match | SDIFF after match | p>\|t\| | SDIFF before match | SDIFF after match | p>\|t\| |
| Mn | -53.3 | -2.9 | 0.768 | -70.2 | -4.9 | 0.358 | -78.2 | -5.4 | 0.516 |
| Opfort_1949 | -34.6 | 2.6 | 0.783 | -14.7 | -1.1 | 0.869 | 48.4 | -0.3 | 0.981 |
| Opfort_1959 | -36.0 | -0.4 | 0.956 | -1.9 | 1.4 | 0.832 | 36.0 | -4.8 | 0.704 |
| Opfort_1976 | 17.3 | -1.3 | 0.914 | 9.5 | -1.1 | 0.877 | -60.5 | -1.2 | 0.841 |
| Opfort_1981 | 34.8 | 3.9 | 0.768 | -14.5 | -0.2 | 0.975 | -35.3 | -1.7 | 0.727 |
| Opfort_1984 | 13.2 | 1.0 | 0.944 | -4.9 | -1.1 | 0.854 | - | - | - |
| Opfort_1997 | 27.5 | -5.7 | 0.669 | -26.0 | 0.5 | 0.896 | - | - | - |

| | Matched A-labelled houses | | | Matched B-labelled houses | | | Matched C-labelled houses | | |
|---|---|---|---|---|---|---|---|---|---|
| # labelled houses | 157 | | | 443 | | | 165 | | |
| # control houses | 712 | | | 712 | | | 712 | | |
| Opfort_1999 | 2.7 | -3.9 | 0.763 | - | - | - | - | - | - |
| Lntotindk | 33.3 | 4.4 | 0.698 | 24.9 | -2.9 | 0.640 | 17.1 | -0.3 | 0.980 |
| Lnboligarl | 42.0 | -2.3 | 0.830 | -4.7 | -3.1 | 0.645 | -67.2 | 0.2 | 0.987 |
| A_yob | -10.0 | -1.2 | 0.909 | -8.9 | -0.6 | 0.922 | -20.2 | -8.6 | 0.385 |
| Antper_a | 32.5 | 4.7 | 0.627 | 16.1 | -1.0 | 0.869 | 6.7 | -0.6 | 0.950 |
| Antper2 | 17.6 | 3.6 | 0.658 | 8.3 | -0.4 | 0.943 | -0.5 | -0.5 | 0.946 |
| Anc_06 | 34.3 | 5.1 | 0.663 | 26.8 | -1.3 | 0.852 | 24.8 | 7.9 | 0.480 |
| Anc_714 | 10.0 | 1.6 | 0.894 | 10.0 | 0.3 | 0.964 | -7.4 | -5.8 | 0.585 |
| Anc_1517_a | -3.1 | -6.5 | 0.574 | 2.4 | -4.3 | 0.544 | 2.1 | -5.7 | 0.627 |
| Lbeton | -30.9 | -1.0 | 0.847 | 14.5 | 3.5 | 0.627 | 27.6 | -0.2 | 0.988 |
| Andet_y | -0.3 | 3.0 | 0.783 | -6.0 | 0.0 | 1.000 | 13.4 | -4.6 | 0.722 |
| Cement | 25.7 | -9.2 | 0.471 | -10.7 | -3.7 | 0.566 | -45.3 | -1.5 | 0.800 |
| Tegl | -10.0 | -0.5 | 0.964 | -5.8 | 2.7 | 0.683 | 33.4 | -3.3 | 0.777 |
| Andet_t | -5.4 | 3.1 | 0.770 | 8.9 | -4.4 | 0.547 | 3.5 | -4.7 | 0.686 |
| Andet_v | 5.6 | 6.3 | 0.571 | -6.7 | -0.3 | 0.957 | -11.4 | 1.9 | 0.820 |
| Andet_o | 2.6 | 5.7 | 0.601 | -1.5 | -2.1 | 0.761 | 10.6 | 4.5 | 0.697 |
| Andet_s | 4.0 | -1.8 | 0.883 | -2.5 | 0.1 | 0.993 | - | - | - |
| Brandovn | -4.0 | 0.5 | 0.963 | 12.6 | -3.7 | 0.601 | -4..6 | -10.2 | 0.365 |
| Pejs | 10.5 | -2.1 | 0.872 | -0.3 | 3.6 | 0.565 | -4.9 | -0.5 | 0.957 |
| Solp | 3.0 | -2.8 | 0.833 | 5.9 | 0.5 | 0.947 | - | - | - |
| Toilet_1 | -53.3 | 1.1 | 0.910 | -0.4 | -1.9 | 0.776 | 36.9 | -1.8 | 0.867 |
| Toilet_3 | 16.8 | -0.2 | 0.986 | 5.6 | -5.2 | 0.487 | 3.0 | -5.8 | 0.630 |
| Bad_2 | 33.7 | -3.8 | 0.751 | -1.4 | 2.7 | 0.689 | -29.8 | -2.0 | 0.841 |
| Bad_3 | 10.6 | -1.9 | 0.894 | 4.5 | -3.4 | 0.668 | - | - | - |
| Etager_2 | 1.7 | 1.8 | 0.873 | 4.1 | -2.8 | 0.704 | 1.0 | -2.1 | 0.855 |
| Niv1 | -18.1 | 1.4 | 0.892 | -20.5 | 3.2 | 0.593 | -12.9 | -1.9 | 0.854 |
| Niv3 | 0.9 | -4.6 | 0.689 | 6.0 | 2.3 | 0.738 | 18.8 | 2.3 | 0.843 |
| Niv4 | 14.2 | 0.7 | 0.957 | 18.0 | 0.9 | 0.905 | 12.1 | -4.1 | 0.732 |
| Niv0 | -4.2 | -3.9 | 0.728 | -10.5 | 1.5 | 0.767 | -0.7 | -1.4 | 0.900 |
| Graddag | 19.8 | 3.5 | 0.755 | 12.1 | 1.7 | 0.799 | 36.8 | -2.4. | 0.825 |

**Table 7 Covariates balance analysis between labelled and non-labelled houses based on Epanechnikov kernel matching. Dependent variable = log (consumption of natural gas / size of house). 4 years after house purchase**

| | Matched A-labelled houses | | | Matched B-labelled houses | | | Matched C-labelled houses | | |
|---|---|---|---|---|---|---|---|---|---|
| # labelled houses | 71 | | | 199 | | | 58 | | |
| # control houses | 366 | | | 366 | | | 366 | | |
| Variable name | SDIFF before match | SDIFF after match | p>\|t\| | SDIFF before match | SDIFF after match | p>\|t\| | SDIFF before match | SDIFF after match | p>\|t\| |
| Mn | -57.1 | -1.0 | 0.946 | -62.8 | -3.7 | 0.653 | -72.8 | -2.2 | 0.880 |
| Opfort_1949 | -40.3 | -2.3 | 0.866 | -16.5 | -7.9 | 0.420 | 57.2 | 5.2 | 0.797 |
| Opfort_1959 | -30.8 | 2.1 | 0.843 | 6.8 | -1.6 | 0.877 | 36.8 | -6.3 | 0.772 |
| Opfort_1976 | 11.8 | -1.1 | 0.951 | 8.7 | -0.6 | 0.956 | -55.9 | 4.7 | 0.675 |
| Opfort_1981 | 30.3 | 7.0 | 0.714 | -6.5 | -0.0 | 0.997 | - | - | - |
| Opfort_1984 | 14.9 | 11.2 | 0.531 | 1.9 | 2.3 | 0.820 | - | - | - |
| Opfort_1997 | 37.5 | -3.1 | 0.880 | -24.6 | -1.5 | 0.830 | - | - | - |
| Opfort_1999 | 8.7 | -2.3 | 0.913 | - | - | - | - | - | - |
| Lntotindk | 48.5 | 8.5 | 0.585 | 25.1 | -2.0 | 0.833 | 15.8 | 0.1 | 0.605 |
| Lnboligarl | 35.1 | -5.0 | 0.757 | -8.2 | -1.7 | 0.858 | -62.4 | -4.8 | 0.800 |
| A_yob | -22.5 | -8.9 | 0.529 | -13.7 | 2.3 | 0.801 | -29.8 | 0.5 | 0.977 |
| Antper_a | 46.3 | 6.3 | 0.671 | 23.6 | 2.3 | 0.786 | 12.4 | 1.2 | 0.939 |
| Antper2 | 25.8 | 4.4 | 0.703 | 8.7 | 2.1 | 0.701 | 2.1 | 0.6 | 0.954 |
| Anc_06 | 40.4 | -2.6 | 0.885 | 23.5 | -0.1 | 0.991 | 11.6 | -2.9 | 0.882 |
| Anc_714 | 18.8 | 8.5 | 0.623 | 6.7 | -0.1 | 0.993 | -13.9 | 7.7 | 0.642 |
| Anc_1517_a | 19.0 | 4.1 | 0.823 | 3.6 | 0.7 | 0.943 | 34.5 | 6.5 | 0.775 |
| Lbeton | -9.2 | 7.0 | 0.614 | 13.8 | 0.1 | 0.994 | 29.8 | -2.3 | 0.916 |
| Andet_y | 0.5 | 4.5 | 0.777 | 1.7 | 4.2 | 0.661 | 25.5 | 10.6 | 0.621 |
| Cement | 44.6 | -4.7 | 0.808 | -10.3 | -8.3 | 0.401 | - | - | - |
| Tegl | -18.9 | 4.6 | 0.761 | 2.8 | -5.5 | 0.594 | 58.2 | 3.1 | 0.879 |
| Andet_t | 7.2 | -2.6 | 0.886 | -2.8 | -6.3 | 0.538 | -12.4 | 0.4 | 0.982 |
| Andet_v | 8.5 | 7.3 | 0.670 | -14.4 | 2.8 | 0.690 | -2.0 | 1.5 | 0.935 |
| Andet_o | -6.0 | -0.0 | 1.000 | -6.5 | 0.2 | 0.983 | 11.4 | -4.9 | 0.813 |
| Andet_s | 12.4 | 8.3 | 0.647 | -0.9 | 3.5 | 0.691 | - | - | - |
| Brandovn | 14.8 | 2.7 | 0.878 | 12.6 | 4.9 | 0.639 | 22.5 | 0.6 | 0.977 |
| Pejs | -13.8 | -3.7 | 0.792 | -17.1 | 0.6 | 0.928 | -0.6 | 10.6 | 0.501 |
| Solp | 12.4 | 10.1 | 0.573 | 3.7 | -16.6 | 0.307 | - | - | - |
| Toilet_1 | -53.1 | 2.3 | 0.875 | 0.1 | -4.1 | 0.681 | 41.0 | -2.7 | 0.884 |
| Toilet_3 | 28.3 | -5.8 | 0.780 | -4.4 | 2.4 | 0.793 | 6.5 | 12.0 | 0.495 |

| | Matched A-labelled houses | | | Matched B-labelled houses | | | Matched C-labelled houses | | |
|---|---|---|---|---|---|---|---|---|---|
| # labelled houses | 71 | | | 199 | | | 58 | | |
| # control houses | 366 | | | 366 | | | 366 | | |
| Variable name | SDIFF before match | SDIFF after match | p>\|t\| | SDIFF before match | SDIFF after match | p>\|t\| | SDIFF before match | SDIFF after match | p>\|t\| |
| Bad_2 | 34.9 | 0.4 | 0.980 | -4.0 | 6.5 | 0.507 | -46.8 | 3.8 | 0.793 |
| Bad_3 | 5.6 | -5.9 | 0.776 | -3.9 | 1.4 | 0.865 | - | - | - |
| Etager_2 | -1.9 | 6.6 | 0.628 | -5.5 | -1.8 | 0.843 | 11.4 | -2.2 | 0.921 |
| Niv1 | -33.6 | -2.3 | 0.868 | -12.2 | 3.7 | 0.693 | -6.7 | 1.2 | 0.948 |
| Niv3 | 9.3 | 1.7 | 0.922 | -9.6 | -1.0 | 0.920 | 20.3 | -4.4 | 0.822 |
| Niv4 | 20.7 | 1.5 | 0.934 | 22.3 | -3.2 | 0.776 | 5.4 | 4.0 | 0.832 |
| Niv0 | -5.8 | -4.5 | 0.783 | -9.4 | 1.3 | 0.862 | -3.3 | -3.8 | 0.842 |

We use the Epanechnikov kernel PSM with bandwidth 0.06.[13] The point and 95% confidence interval estimated ATT for each sub-sample are presented in table 8. Columns 3-5 show the estimated ATT for A-labelled houses in 1-4 years after the house is bought, columns 7-9 present the estimated ATT for B-labelled houses in 1-4 years after the house is bought, and columns 11-13 present the estimated ATT for C-labelled houses in 1-4 years after the house is bought.

**Table 8 Propensity Score Matching Estimation of Average Treatment Effect for the Treated**

| | A-labelled house | | | |
|---|---|---|---|---|
| Years after house purchase | $N_1$ | [ | ~ ATT | ] |
| 1 | 1362 | -0.214 | -0.141 | -0.103 |
| 2 | 1417 | -0.129 | -0.082 | -0.023 |
| 3 | 869 | -0.078 | -0.009 | 0.077 |
| 4 | 437 | -0.212 | -0.068 | 0.042 |
| | B-labelled house | | | |
| Years after house purchase | $N_1$ | [ | ~ ATT | ] |
| 1 | 2018 | -0.063 | -0.009 | 0.023 |
| 2 | 1889 | -0.032 | -0.001 | 0.060 |
| 3 | 1155 | -0.024 | 0.012 | 0.081 |
| 4 | 565 | -0.111 | -0.059 | 0.006 |
| | C-labelled house | | | |

[13] The psmatch2 procedure is used in this application (see Leuven and Sianesi 2003).

| | A-labelled house | | | |
|---|---|---|---|---|
| Years after house purchase | $N_1$ | [ | ~ ATT | ] |
| Years after house purchase | $N_1$ | [ | ~ ATT | ] |
| 1 | 1369 | -0.027 | 0.073 | 0.229 |
| 2 | 1422 | -0.073 | -0.010 | 0.072 |
| 3 | 877 | -0.072 | 0.007 | 0.078 |
| 4 | 424 | -0.115 | 0.032 | 0.169 |

Note: [ = 95% lower bound. ] = 95% upper bound. ATT: point estimate of Average Treatment Effect.

Epanechnikov kernel PSM with bandwidth 0.06. Bootstrap Standard errors used to construct confidence interval (see Leuven and Sianesi.2003)[14]

Practically all ATT, with the exception of A-labelled houses for years 1 and 2 after labelling, are insignificant at 5%, rejecting the hypothesis of average energy savings due to labelling in Denmark several years after a house is labelled. We find only a significant negative ATT for A-houses for years 1 and 2. However, the energy saving effect for the first year is bigger in absolute value than the saving effect for the second year, a somehow surprising result that in any case supports the expected saving pattern of an effective energy labelling scheme. This result might indicate the presence of selection bias because of non-observable heterogeneity at A-labelled houses.

Sensitivity analyses have been conducted to test for the robustness of the results. The results are robust to different specifications for the PSM algorithm. Concretely, estimated ATT with a logit propensity score, different bandwidths, or bi-weight kernel do not depart from the reported results in the paper.[15] Summing up, our empirical results do not support any of the hypotheses associated with an effective energy labelling.

A possible explanation for the non-significant reductions in consumption could be that especially the C-labelled houses are significantly older (see table 3), and it is therefore likely that many other things like kitchen and bathroom need to be updated before a new owner prioritises to improve the energy efficiency. Another explanation that covers all the labelled houses can be found in a study made by Jensen (2004), where it is investigated why house owners do not invest in energy efficient solutions. Among other things he finds that it is not classical barriers like money constraints, lack of interest or knowledge. According to him the problem is that house owners find other factors more important than consumption of energy and energy saving. For them the visual improvements of the house are more important, e.g. new kitchen, new bathroom. A study conducted by Danish Energy reveals another explanation. It shows that only 20% of house owners are willing to spend more than 30,000 DKK on energy improvements of their house, and they are not willing to accept a return on their investment on more than 6 years (see Danish Energy 2007).

The Madsen, Ramlau and Pedersen (2001) evaluation finds that 30% of the people in their interview survey are willing to accept a return on their investments on 5-7 years. 21% are only willing to accept a full return within 3-4 years, while 13% are willing to wait 8-11 years for a return of their investment. Unfortunately, a large amount of the proposed invest-

14 Results of other bandwidths are available upon request.

15 Results from the sensitivity analysis can be presented upon request.

ments is not paid back before 20-40 years after the investment is made, and if the respondents really act according to their answers in the interview survey, these proposed investments in energy saving improvements will never be carried out (see Madsen, Ramlau and Pedersen 2001).

Both studies point at the importance of the pay-back period for the investment being rather short. This is a well known problem, also used as argument in papers on the presence of energy efficiency gap. Here the argument is that energy efficiency gaps occur when the discount factor used for calculating the present value of the investment minus energy savings is not the same as the discount factor, house owners act according to. House owners and consumers in general are impatient and will rather spend money now, than wait for a saving occurring after several years; hence they have a high discount factor.

# 6 Conclusion

The main purpose of this paper has been to evaluate the Danish Energy Labelling Scheme. The evaluation has been carried out by merging data on consumption of natural gas for heating in single-family houses with register data on house specific characteristics and household characteristics. We have used propensity score matching to estimate average treatment on the treated (ATT) – average labelling effect in terms of natural gas consumption on labelled houses. Propensity score matching has been carried out on 12 sub-samples; A-, B- and C-labelled houses observed 1-4 years after the house purchase, which is also the time, where the new house owners get the energy label information. The sub-samples were constructed to examine whether the effect of the energy labelling scheme on energy consumption – if such could be found significant – would depend on the energy related state of the house and/or would be related to the time passed since the house was labelled. The hypothesis being: if a significant effect on consumption of natural gas could be found, then we would expect it to be negative, and we would expect the greatest effect for C-labelled houses, since they are in worst conditions, and would therefore be subject to the largest number of recommended improvements or the largest investments related to such improvements.

With exception of A-labelled houses in the first two years after house purchase, all the estimated ATT are insignificant, and therefore our empirical results cannot support the hypothesis of significant average energy savings due to the Danish Energy Labelling Scheme several years after a house is bought. Even the significant result for the A-labelled houses cannot support the expected saving pattern of an effective energy labelling scheme, since the estimated ATT – energy saving – is bigger the first year than the second year.

The empirical results of this paper support the findings from other studies on the Danish Energy Labelling Scheme. For instance, the Madsen, Ramlau and Pedersen (2001) study that found very small and close to statistical insignificant differences in investment levels and energy savings in labelled and non-labelled houses, Jensen (2004) who identified a number of barriers for realising energy savings in buildings and the Gram-Hanssen and Jensen (2006) study that found that the respondents remembered the label, they found that the labelling scheme was a good idea, but they have not really used the information from the label.

Even though evaluating this data could not reveal any significant reduction in natural gas consumption for energy labelled houses compared to non-labelled houses, the labelling scheme might still have effect. If the owner of a labelled house implements some or all of the recommended improvements and thereby can obtain the same indoor temperature at a lower energy level/price, he might decide to raise the indoor temperature with the saved amount of energy, and reach a higher utility level, instead of saving the money and energy. This would be a welfare gain, but since indoor temperature is not registered, it would not be possible to use this change for estimation of the labelling effect.

# 7 References

Building Energy Labelling Forum Agenda (2006): *Building Energy Labelling Forum* held Friday December 1, 2006, in Toronto. Office of Energy Efficiency, Commercial and Institutional Buildings, Natural Resources Canada.

Danish Energy (2007): *CATINÉT research for ELFOR March 2007*.

Dehejia, R.H. and S. Wahba (1999): Causal Effects in Non-Experimental Studies: Evaluating the Evaluation of Training Programs. *Journal of the American Statistical Association* 94: 1053-1062.

Ditlefsen, J. (2008): *Email correspondence*. December 2008.

Gram-Hanssen, K. and O.M. Jensen (2006): *Energimærkning af enfamiliehuse – en kvalitativ analyse*. Notat. Statens Byggeforskningsinstitut [Danish Building Research Institute]. Only available in Danish.

Heckman, J.J. and R. Robb (1985): Alternative Methods for Evaluating the Impact of Interventions. In: J.J. Heckman and B. Singer (eds.): *Longitudinal Analysis of Labor Market Data*. New York, Cambridge University Press: 156-245.

Heckman, J.J., H. Ichimura and P. Todd (1997): Matching as an Econometric Evaluation Estimator. *Review of Economic Studies* 65: 261-294.

Jensen, O.M. (2004): *Barrierer for realisering af energibesparelser i bygninger. Statens Byggeforskningsinstitut*. Only available in Danish.

Laustsen, J. and K. Lorenzen (2003): *Danish Experience in Energy Labelling of Buildings*. COWI and Danish Energy Authority.

Leuven, E. and B. Sianesi (2003): psmatch2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing.

Madsen, B.B.; M. Ramlau and N.B. Pedersen (2001): *Evaluering af Energimærkningsordningen. Slutrapport*. Made by COWI for Energistyrelsen. Only available in Danish.

OECD (2003): *Environmentally Sustainable Buildings. Challenges and Policies*. OECD.

Petersen, K.N. and K. Gram-Hanssen (2005): *Husholdningernes energi- og vandforbrug. Afhængighed af socioøkonomiske baggrundsvariable*. SBI 2005:09. Statens Byggeforskningsinstitut. Only available in Danish.

RESNET homepage (2008): www.natresnet.org. Homepage for Residential Energy Network Service.

Rosenbaum, P.R. and D.B. Rubin (1983): The Central Role of Propensity Score Matching in Observational Studies for Causal Effects. *Biometrika* 70: 41-55.

Rosenbaum, P. and D. Rubin (1985): Reducing Bias in Observational Studies Using Subclassification on the Propensity Score. *Journal of the American Statistical Association* 79: 516-524.

Rubin, D.B. (1974): Estimating Causal Effects of Treatment in Randomized and Nonrandomized Studies. *Journal of Education Psychology* 66: 688-701.

Thomsen, K.E.; K.B. Wittchen, O.M. Jensen and S. Aggerholm (2006): *Applying the EPBD to improve the Energy Performance Requirements for Existing Buildings – ENPER-EXIST*. WP3: Building stock knowledge.

AKF

Nyropsgade 37
DK-1602 København V

tel: +45 4333 3400
fax: +45 4333 3401

akf@akf.dk
www.akf.dk