# Energy efficiency premium issues and revealing the pure label effect

Aras Khazal [a*], Ole Jakob Sønstebø [a]

[a] *NTNU Business School, Norwegian University of Science and Technology, Trondheim, Norway*

[*] *Corresponding author. E-mail address:* *aras.khazal@ntnu.no*

## Abstract

Following the European Union's implementation of Energy Performance Certificates (EPCs) for buildings, the capitalization of energy efficiency in transaction prices and rents has been subject to much research. This paper uses different identification strategies for the Norwegian residential sales ($N$ = 750,000) and rental ($N$ = 670,000) markets to highlight the endogeneity and methodological limitations associated with assessing the price effects of energy efficiency and the signaling effect of label adoption. We find that the valuation of energy efficiency is subject to unobserved location and quality bias, that labeling has immediate, short-run, and long-run price effects and that different effects are observed in different submarkets. We provide evidence that sample selection issues related to location and time, with methodological and data limitations, are essential factors that must be considered when assessing the effects of the EPC implementation.

***Keywords:*** Energy performance certificates, residential rental market, residential sales market, environment, housing policy

JEL: C01; P25; Q4; R10; R20

# 1 Introduction

Climate change continues to be one of the most topical concerns of political discussion and decision-making. Measures have been taken in various parts of commercial trade to reduce energy consumption and carbon emission. The energy efficiency of buildings is a key factor for increased environmental sustainability in the housing sector, and certain regulatory actions have been made in this regard. In the EU, Energy Performance Certificates (EPCs) were introduced due to the Energy Performance of Buildings Directive in 2002. The EPC is a legal document that summarizes the energy efficiency of buildings, rated from A (low expected energy consumption) to G (high expected energy consumption). The EPC policy is aimed to provide information and awareness regarding the energy efficiency of buildings and to create economic incentives for actors to invest in environmentally friendly improvements of dwellings, since improvement of a dwelling's energy performance is expected to yield higher transaction prices and rents (Mudgal et al., 2013, p. 18). Through the CONSEED (CONSumer Energy Efficiency Decision making) project, funded by Horizon 2020, the biggest EU Research and Innovation programme to date, evidence is found that consumers are willing to pay more for energy efficient properties when the energy costs are disclosed. Research through the ongoing Horizon 2020 funded PENNY (Psychological, social and financial barriers to energy efficiency) project sets out to analyze behavioral factors, purchasing decisions, institutional conditions and broader implications regarding energy efficiency and policies in order to support the development of energy efficiency strategies, policies and programmes across Europe.

The capitalization of energy efficiency in transaction prices and rents has been subject to much research in recent years. For example, in commercial office markets, Eichholtz et al. (2010), Wiley et al. (2010), Fuerst and McAllister (2011a), and Kok and Jennen (2012) found evidence of capitalization in both rental and sales prices, whereas Fuerst and McAllister (2011b) concluded that no relation exists between EPCs and rental prices. In residential transaction markets, Murphy

(2014), Wahlström (2016), Hårsman et al. (2016), and Fregonara et al. (2017) discovered no evidence of capitalization, whereas Brounen and Kok (2011), Ramos et al. (2015), de Ayala et al. (2016), Davis et al. (2015), Wilhelmsson (2019), Jensen et al. (2016), and Högberg (2013) found a premium for more efficient dwellings. Aydin et al. (2017) also discovered that energy efficiency is associated with higher prices but that the signaling effect of the EPCs is not significant, as opposed to Jensen et al. (2016) who find the signaling effect to be significant. In residential rental markets, Salvi et al. (2010), Hyland et al. (2013), Cajias and Piazolo (2013), and Kholodilin et al. (2017) indicated that energy-efficient homes are associated with higher rents compared with inefficient homes. A small number of existing EPC studies exist on the Norwegian real estate market. In the residential sales market, Olaussen et al. (2017) and Olaussen et al. (2019) discovered no evidence that the disclosure of energy efficiency has any effect on prices, whereas Khazal and Sønstebø (2020) discovered a positive effect on residential rental prices. Table A1 in the Appendix offers an overview of the EPC literature with methods and major findings.

It is challenging to assess the price implications of the EPC policy implementation. Confounding factors for the impact on prices could include sample selection (size, location, and time window) and endogeneity problems originating from omitted variables correlated with energy efficiency. Only a few studies thoroughly address this issue; for example, Wilhelmsson (2019) applied propensity score matching, spatial dependency, and controls for outliers and selection bias, and Aydin et al. (2017) applied instrumental variable (IV) and repeat sales approaches to overcome potential bias. Hence, it is still necessary to investigate the EPC policy effects on prices using different strategies: both to identify confounding factors behind the price impact of energy efficiency and to reveal the pure price effect of disclosing the dwelling's energy efficiency through label adoption.

In this paper, we apply different identification strategies using highly representative samples for both the rental and sales markets. First, we thoroughly control for the locational effects by employing pooled cross-sectional high-dimensional fixed effects (HDFE) and panel fixed effects (FE) approaches. Next, we apply the IV approach using the total number of new EPCs (NNC) as an instrument to assess the pure label effect on prices of disclosing the dwelling's energy efficiency through label adoption. Finally, we investigate the signaling impact of label adoption over time by recursively estimating the IV regression over the sample periods.

We find that the valuation of energy efficiency is subject to endogeneity originating from unobserved locational factors, and that dwellings with lower energy efficiency are associated with more locational bias in the rental market, while this bias is higher for the energy efficient dwellings in the sales market. Further, we find that the lower the energy efficiency, the less bias comes from unobserved quality in the sales market. Overall, improving the energy efficiency of the dwelling with one letter on the EPC rating has similar effects for both rental and sales objects, with a price impact of about 0.8-1.0%.

Investigating the signaling effect of label adoption, we discover that the IV approach produces biased estimations if the location is not controlled for. However, after considering this issue, both the IV and panel FE demonstrate that a positive signaling effect exists on prices in both markets. In the rental market, the IV approach provides evidence that the main source of endogeneity is due to locational heterogeneity rather than unobserved quality. The recursive estimations reveal that labeling has immediate, short-run, and long-run price effects and that different effects are observed in different submarkets. The findings also highlight the possibility that different conclusions might be drawn due to sample selection issues related to time periods and submarkets, and that methodological and data limitations are essential factors that must be considered when assessing the effects of the EPC implementation.

The rest of the paper is structured as follows. In the next section, we offer a description of the rental and sales data used in the analyses. Section 3 comprises the methodology and analyses, and Section 4 concludes.

## 2 Data and Descriptive Statistics

In Norway, the EPC policy was implemented on July 1, 2010, and the certificate follows the EU standard. At any time, the document can be updated with new information submitted by the homeowner or landlord, and the certificate is valid for 10 years. Although the EU Commission requires that certificates be included in all advertisements in commercial media when a building is announced for sale or rent, a majority of dwellings are still not labeled in the Norwegian rental market. The reasons for the insufficient labeling by landlords could be the diffusion of information from the managing body and the inefficient penalty system (Khazal and Sønstebø, 2020). However, in the Norwegian sales market, where the pricing mechanism is characterized as an English auction, most dwellings are labeled.

Hyland et al. (2016) find evidence of bunching in the Irish residential property market, where energy efficiency assessors tend to over-evaluate the energy performances of dwellings at the thresholds between labels – indicating a potential measurement error in the labels. The authors speculate that this happens because the assessors want to generate repeat business and improve their reputation. However, Norwegian homeowners and landlords have the option of cost-free self-assessment when obtaining an EPC, and the certificate is automatically produced based on the information input. Hence, there is no third-party agent with hidden intentions and little incentive for the behavior found in Ireland. The certification system is trust-based, but it is illegal to provide wrongful information and the homeowner has the legal responsibility. Additionally, insurance is often bought to ensure that all information regarding the dwelling, including the EPC, is correct. Hence, measurement error in the EPCs seems to be less likely in the Norwegian market.

## 2.1 The rental market data

In this study, we use two datasets for residential dwellings, presented in Tables 1 and 2, both provided by Norway's largest and most frequently used online advertisement site, Finn.no. The first dataset comprises nearly 670,000 rental advertisements over the period from 2011 to 2019, containing posted monthly rents, EPC information, dwelling characteristics, several property amenities, and location at the county, municipality, and zip code levels. Because we are unable to follow each unit over time, the rental data are treated as pooled cross-sectional data. The average rent for the sample period is EUR 1,031, and about 16% are labeled dwellings, distributed somewhat equally among the specific labels. The dataset is highly representative, comprising observations from all counties, 98% of municipalities, and 74% of zip codes in Norway. Boligbygg Oslo KF is a Norwegian municipal corporation for housing that every quarter, since the autumn of 2005, have conducted a separate survey aimed at those who have advertised their apartment for rent in the last quarter. They are able to link the answers from the survey with the information in the ad. In the survey they ask a number of questions that provides the opportunity to correct the difference between the advertised and actual (contract) rental price. The results show that the differences are very small (up to 1 percent) on average (Skogstad, 2011). Full descriptive statistics on the rental data are reported in Table 1.

Table 1: Summary statistics of the rental data

| Variable | Mean | SD |
|---|---|---|
| Rental price (€) | 1,032 | 406 |
| Energy label (*percent*) | | |
| A | 02.78 | 16.45 |
| B | 02.21 | 14.70 |
| C | 02.76 | 16.38 |
| D | 02.50 | 15.60 |
| E | 01.67 | 12.81 |
| F | 01.37 | 11.61 |
| G | 02.93 | 16.87 |
| Non-labeled | 83.78 | 36.86 |
| Floor location | 1.24 | 1.60 |
| Size ($m^2$) | 65.34 | 31.68 |
| Bedrooms (*percent*) | 1.69 | 0.99 |
| Age information (*percent*) | 00.65 | 08.06 |
| Furnished (*percent*) | 17.42 | 37.93 |
| Centrally located (*percent*) | 46.43 | 49.87 |
| Nice view (*percent*) | 25.04 | 43.32 |
| Fireplace (*percent*) | 17.44 | 37.95 |
| Janitor service included (*percent*) | 12.30 | 32.85 |
| Hiking accessibility (*percent*) | 37.89 | 48.51 |
| Security alarm (*percent*) | 04.97 | 21.72 |
| Nonsmoking (*percent*) | 00.02 | 01.45 |
| Public service accessibility (*percent*) | 00.23 | 04.75 |
| Quiet area (*percent*) | 46.23 | 49.86 |
| Garden (*percent*) | 00.56 | 07.49 |
| Dwelling type (*percent*) | | |
| Detached | 08.34 | 27.65 |
| Bedsit | 11.23 | 31.57 |
| Apartment | 75.85 | 42.80 |
| Townhouse | 01.32 | 11.41 |
| Semi-detached | 03.26 | 17.76 |
| Year of transaction (*percent*) | | |
| 2011 | 09.24 | 28.96 |
| 2012 | 09.99 | 29.99 |
| 2013 | 11.01 | 31.31 |
| 2014 | 13.07 | 33.71 |
| 2015 | 15.15 | 35.85 |
| 2016 | 14.10 | 34.80 |
| 2017 | 10.95 | 31.22 |
| 2018 | 09.46 | 29.26 |
| 2019 | 07.03 | 25.56 |
| Quarter of transaction (*percent*) | | |
| $Q_1$ | 24.22 | 42.84 |
| $Q_2$ | 26.82 | 44.30 |
| $Q_3$ | 28.23 | 45.01 |
| $Q_4$ | 20.73 | 40.53 |
| No. of counties *(% of Norway)* | 18 (100%) | |
| No. of municipalities *(% of Norway)* | 417 (98%) | |
| No. of zip codes *(% of Norway)* | 3583 (74%) | |

Note: Table 1 reports the summary statistics of the Norwegian rental data. The dataset comprises nearly 670,000 rental advertisements over the period from January 2011 to September 2019. Data source: Finn.no.

## 2.2 The sales market data

The second dataset comprises nearly 750,000 sales advertisements over the period from 2010 to 2017, containing asking prices, EPC information, dwelling characteristics, and location. The advantage of this dataset is the inclusion of the street address information, which allows us to treat the data as both pooled cross-sectional data and as a panel set with 71,966 multiple-sale units over the sample period ($N$ = 177,634). The average asking price is about EUR 286,000. Moreover, 63% of dwellings are labeled, and the proportions of the labels are decreasing with increased energy efficiency. The data contain information about dwellings located in all counties, all municipalities, and about 73% of the Norwegian zip codes. Homes in Norway are sold in a perfectly competitive bidding context, where the potential buyers compete with open bids and the highest wins the auction. The Norwegian market laws prohibit intentionally listing a dwelling below the expected market value in order to attract more buyers (see for example Han and Strange, 2016). Hence, the asking price is expected to reflect the true market value. Lyons (2019) investigates whether listings accurately capture sales price trends in the Irish housing market over the period 2006-2012 and concludes that listing prices are an adequate substitute in cases where transaction data are not available. In this paper, we use asking prices as a proxy for sales prices and refer to prices as sales prices hereafter. Table 2 lists the full descriptive statistics of the sales data.

Table 2: Summary statistics of the sales data

| Variable | Mean | SD |
|---|---|---|
| Asking price (€) | 286,131 | 195,952 |
| Common debt (€) | 10,806 | 29,715 |
| Size ($m^2$) | 104.85 | 58.68 |
| Dwelling age (*years*) | 43.14 | 33.84 |
| Bedrooms (*percent*) | 2.57 | 1.23 |
| Energy label (*percent*) | | |
| A | 0.20 | 04.51 |
| B | 1.88 | 13.59 |
| C | 5.74 | 23.26 |
| D | 11.59 | 32.01 |
| E | 10.85 | 31.10 |
| F | 14.48 | 35.19 |
| G | 18.28 | 38.65 |
| Non-labeled | 36.99 | 48.28 |
| Dwelling type (*percent*) | | |
| Multi-family house | 00.04 | 02.11 |
| Detached | 31.03 | 46.26 |
| Leisure home | 00.55 | 07.38 |
| Apartment | 55.64 | 49.68 |
| Townhouse | 06.78 | 25.14 |
| Semi-detached | 05.95 | 23.66 |
| Ownership type (*percent*) | | |
| Stock | 01.96 | 13.86 |
| Coop | 29.71 | 45.70 |
| Freehold | 68.34 | 46.52 |
| Year of transaction (*percent*) | | |
| 2010 | 12.36 | 32.92 |
| 2011 | 12.91 | 33.53 |
| 2012 | 12.79 | 33.40 |
| 2013 | 12.75 | 33.35 |
| 2014 | 12.14 | 32.66 |
| 2015 | 12.43 | 32.99 |
| 2016 | 11.83 | 32.30 |
| 2017 | 12.79 | 33.40 |
| Quarter of transaction (*percent*) | | |
| $Q_1$ | 23.15 | 42.18 |
| $Q_2$ | 31.78 | 46.56 |
| $Q_3$ | 25.46 | 43.56 |
| $Q_4$ | 19.61 | 39.71 |
| No. of counties *(% of Norway)* | 18 (100%) | |
| No. of municipalities *(% of Norway)* | 423 (100%) | |
| No. of zip codes *(% of Norway)* | 3,568 (73%) | |

Note: Table 2 reports the summary statistics of the Norwegian sales data. The dataset comprises nearly 750,000 sales advertisements over the period from January 2010 to December 2017. Data source: Finn.no.

# 3 Methodology and Results

We take a unique approach to studying potential endogeneity issues related to the EPC valuation in the residential rental and sales markets. First, we investigate the potential bias originating from unobserved locational heterogeneity by applying the HDFE method on the cross-sectional data and panel FE on the panel dataset. Next, we study the potential bias related to unobserved quality characteristics applying the IV approach to reveal the pure label effect on prices.

## 3.1 Energy efficiency premium issues

The most common approach for modeling housing rents and prices is the hedonic model developed by Court (1939) and Rosen (1974), which expresses the price as a function of a set of dwelling characteristics. Our datasets have multiple geographical levels that each may have attributes correlated with energy performance. For example, the more populous cities have a more clustered building structure, which could be associated with higher energy efficiency and consequently with a higher EPC rating. The overall quality, weather conditions, and other locational factors of a neighborhood might also be correlated with energy performance. However, such locational control variables are often difficult to obtain, and sometimes unobservable. The omission of one or more of these variables invalidates the ordinary least squares (OLS) assumption that the EPCs are independent of the error term and lead to biased estimates. Hence, the use of small locational entities (such as postcodes or exact dwelling addresses) is preferable to the use of larger ones with respect to appropriately controlling for omitted variable bias (Von Graevenitz and Panduro, 2015). To address this issue, we use the HDFE method, developed by Guimaraes and Portugal (2010) and Gaure (2013), among others, which is appropriate for both pooled cross-sectional and panel data. For the pooled cross-sectional data, the hedonic HDFE model can be expressed as follows:

$$ln\, price_{icmz} = X_{icmz}\,\gamma \ + \alpha_c + \mu_m + \omega_z + \lambda_t + \varepsilon_{icmz}. \qquad (1)$$

Here, $ln\, price_{icmz}$ represents the natural logarithm of the rental price or sales price for dwelling $i (i = 1, \ldots, N)$ located at county $c$ $(c = 1, \ldots, C)$, municipality $m$ $(m = 1, \ldots, M)$, and zip code

$z(z = 1, \ldots, Z)$. Moreover, $X_{icmz}$ is a vector of hedonic and other dwelling characteristics (Tables 1 and 2). In addition, $X_{icmz}$ also includes independent dummy variables for each EPC label ranging from A to G, where the default is the non-labeled dwellings. Furthermore, $\gamma$ is the corresponding set of coefficients to be estimated using least squares. All time-invariant characteristics of county, municipality, and zip codes are captured by the FE $\alpha_c$, $\mu_m$, and $\omega_z$, respectively, whereas $\lambda_t$ is a vector of the time FE (year, quarter, and month) and $\varepsilon_{icmz}$ is the error term that is assumed to be uncorrelated with all independent variables. For the panel sales data, the hedonic FE model is specified as follows:

$$ln\, price_{it} = X_{it}\, \gamma\ + \theta_i\ + \lambda_t + \varepsilon_{it}, \qquad (2)$$

where $price_{it}$ is the sales price for dwelling $i$ observed at time $t$, and $\theta_i$ represents the entity (dwelling) FE, which model the effects of omitted time-invariant dwelling heterogeneity on the $price_{it}$. The appropriateness of implementing the panel FE method to estimate hedonic equations relies on whether the variables of interest vary over time. In general, hedonic dwelling characteristics, such as the size and number of bedrooms, are time-invariant and therefore ignored in the model. Similarly, the EPC variable is omitted if the label remains in the same category for each sale throughout the sample period. Hence, the panel FE model only applies to dwellings where the EPC label changes over time, e.g., from non-labeled to C or from F to E. Note that time-variant heterogeneity, such as quality improvements and refurbishments, is not captured by the FE. The inclusion of time FE helps in this regard but cannot rule out the issue completely.

### 3.1.1 Rental market results

Table 3 reports six specifications of equation (1) for the cross-sectional rental data estimations. We start by ignoring all locational effects in column 1 (OLS), and in the next specifications, we estimate the HDFE model by gradually adding locational effects for the county, municipality, and zip code to identify how the coefficient changes when controlling for a higher proportion of locational variation. The OLS estimation in column 1 reveals that the coefficients of energy efficiency are positive and significant compared to non-labeled dwellings. However, the results indicate that less efficient dwellings are associated with higher premiums than more efficient dwellings, when location is not controlled for. Gradually controlling for location in columns 2-4, shows that premiums are subject to endogeneity originating from location. While the premiums of the most efficient dwellings, A and B, are the least affected but increasing slightly with increased location control, the opposite is observed among the least efficient dwellings. Additionally, we re-estimate equation (1) with fully locational control including a continuous variable for energy efficiency ranging from non-labeled to A (column 5) and ranging from G to A excluding the non-labeled dwellings (column 6). The results show that, ceteris paribus, an increase in energy efficiency by one letter yields a premium of about 0.9-1.0% on average. Combined with the substantial increase in $R^2$ and decrease in Root Mean Square Error (RMSE), the results reveal that the effect of energy efficiency on rents is positively biased when not sufficiently controlling for unobserved locational effects.

Further, we apply the Heckman (1979) two-step selection model to control for the possibility that labels may be subject to self-selection bias. In the first step Probit estimation, we used the share of vote for the Green Party at the municipality level from all municipal, county and state elections over the period 2009-2019 as an exogenous determinant of label adoption. In the second step, the coefficient of the selection variable (the inverse Mills ratio) is negative and significant, indicating that unobserved factors behind the adoption of labeling is associated with lower rental prices. However, the coefficients of the energy efficiency premiums are robust even after controlling for the potential sample selection bias, reported in Table A2 in the Appendix.

Table 3: Cross-sectional high-dimensional fixed effects (HDFE) estimations of rental prices

| | OLS (1) | HDFE (2) | HDFE (3) | HDFE (4) | HDFE (4) | HDFE (6) |
|---|---|---|---|---|---|---|
| Efficiency | | | | | 0.0096*** | 0.00877*** |
| | | | | | (71.75) | (26.62) |
| A | 0.0544*** | 0.0578*** | 0.0621*** | 0.0622*** | | |
| | (28.83) | (36.70) | (44.17) | (46.26) | | |
| B | 0.0485*** | 0.0549*** | 0.0562*** | 0.0565*** | | |
| | (22.36) | (30.84) | (36.15) | (38.26) | | |
| C | 0.0637*** | 0.0612*** | 0.0612*** | 0.0523*** | | |
| | (31.90) | (38.59) | (44.42) | (40.04) | | |
| D | 0.0630*** | 0.0503*** | 0.0424*** | 0.0383*** | | |
| | (29.55) | (29.34) | (28.92) | (27.65) | | |
| E | 0.0597*** | 0.0280*** | 0.0153*** | 0.0159*** | | |
| | (22.88) | (14.00) | (8.98) | (9.81) | | |
| F | 0.0930*** | 0.0289*** | 0.0119*** | 0.0109*** | | |
| | (31.46) | (13.14) | (6.10) | (5.93) | | |
| G | 0.150*** | 0.0477*** | 0.0310*** | 0.0226*** | | |
| | (74.80) | (30.86) | (22.62) | (17.54) | | |
| Reference | Non-labeled | Non-labeled | Non-labeled | Non-labeled | | |
| Controls | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| County | ☓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Municipality | ☓ | ☓ | ✓ | ✓ | ✓ | ✓ |
| Zip code | ☓ | ☓ | ☓ | ✓ | ✓ | ✓ |
| Observations | 669,894 | 669,825 | 669,817 | 669,448 | 669,448 | 108,276 |
| Adjusted $R^2$ | 0.405 | 0.614 | 0.702 | 0.737 | 0.725 | 0.746 |
| RMSE | 0.285 | 0.229 | 0.202 | 0.189 | 0.194 | 0.170 |

Note: Table 3 reports the OLS and HDFE estimation results for the rental data. The dependent variable is the natural logarithm of the monthly rental price. The default for EPC-labels is non-labeled dwellings. Control variables stand for hedonic characteristics (Table 1) and time fixed effects (year, quarter, and month). Fixed location effects for county, municipality, and zip code. The sample period spans from January 2011 to September 2019. The subsample of labeled-only dwellings is used in HDFE (6). Heteroskedasticity robust *t*-statistics are in parentheses. * $p < .05$, ** $p < .01$, *** $p < .001$.

Additionally, we run several regressions based on different stratifications of the data, reported in Table 4. The energy efficiency premium of the largest cities is about 0.8% compared to 1.0% for the rest of Norway, and the results also indicate that the premium for detached homes is slightly higher than for apartments and other types of dwellings. We also observe that the premium is increasing in dwelling size. Nevertheless, in general, the premium is not substantially deviating from the premium reported in Table 3, columns 5 and 6.

Table 4: Heterogeneity analyses of the impact of energy efficiency on rental prices

| | Largest cities | Largest cities excluded | Apartment | Detached | Other types | Small $\leq 60$ m$^2$ | Medium $> 60 \leq$ 120 m$^2$ | Large $> 120$ m$^2$ |
|---|---|---|---|---|---|---|---|---|
| Efficiency | 0.0082*** | 0.0099*** | 0.0090*** | 0.0108*** | 0.0093*** | 0.0078*** | 0.0099*** | 0.0119*** |
| | (41.52) | (57.05) | (65.23) | (16.30) | (70.95) | (50.58) | (46.54) | (14.15) |
| Controls | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Observations | 342,649 | 326,799 | 507,717 | 55,518 | 669,448 | 360,851 | 266,632 | 41,314 |
| Adjusted $R^2$ | 0.692 | 0.702 | 0.730 | 0.686 | 0.7364 | 0.743 | 0.679 | 0.625 |
| RMSE | 0.193 | 0.184 | 0.178 | 0.229 | 0.190 | 0.161 | 0.199 | 0.258 |

Note: The dependent variable is the natural logarithm of the rental price. Controls stand for hedonic characteristics (Table 1), year, quarter, and month fixed effects (FE) and county, municipality, and zip codes FE. The sample period is from January 2011 to September 2019. The largest Norwegian cities are defined as having a population exceeding 90,000 (Oslo, Bergen, Stavanger, Sandnes, Trondheim, Fredrikstad, Sarpsborg, Drammen, and Porsgrunn). Heteroskedasticity robust *t*-statistics are in parentheses. $^{*}$ $p < .05$, $^{**}$ $p < .01$, $^{***}$ $p < .001$.

### 3.1.2 Sales market results

In Table 5, we apply the same identification strategy for the sales data to investigate the potential locational bias by estimating equation (1) in columns 1-4 and equation (2) in columns 5-8. All energy efficiency coefficients are positive and significant and decrease in magnitude from the OLS through the HDFE specifications (columns 1-4), when gradually including county, municipality, and zip code FE.

Table 5: Cross-sectional high-dimensional fixed effects (HDFE) and panel FE estimations of sales prices

| | OLS (1) | HDFE (2) | HDFE (3) | HDFE (4) | Panel FE (5) | Panel FE (6) | Panel FE (7) | Panel FE (8) |
|---|---|---|---|---|---|---|---|---|
| Efficiency | | | | | | | $0.0079^{***}$ | $0.0085^{***}$ |
| | | | | | | | (20.20) | (9.49) |
| A | $0.1911^{***}$ | $0.1501^{***}$ | $0.1670^{***}$ | $0.1700^{***}$ | $0.0636^{***}$ | $0.0743^{**}$ | | |
| | (14.29) | (12.88) | (15.78) | (17.25) | (3.35) | (2.90) | | |
| B | $0.1967^{***}$ | $0.1590^{***}$ | $0.1469^{***}$ | $0.1286^{***}$ | $0.0225^{***}$ | $0.0181^{*}$ | | |
| | (53.88) | (50.80) | (53.82) | (51.71) | (3.43) | (2.24) | | |
| C | $0.1756^{***}$ | $0.1384^{***}$ | $0.1207^{***}$ | $0.1024^{***}$ | $0.0357^{***}$ | $0.0352^{***}$ | | |
| | (71.10) | (64.68) | (64.94) | (61.26) | (11.00) | (9.60) | | |
| D | $0.1106^{***}$ | $0.0872^{***}$ | $0.0660^{***}$ | $0.0581^{***}$ | $0.0336^{***}$ | $0.0356^{***}$ | | |
| | (60.77) | (55.90) | (48.93) | (48.44) | (16.03) | (15.06) | | |
| E | $0.0315^{***}$ | $0.0237^{***}$ | $0.0028^{*}$ | $0.0156^{***}$ | $0.0248^{***}$ | $0.0284^{***}$ | | |
| | (18.17) | (16.43) | (2.38) | (15.37) | (14.16) | (14.45) | | |
| F | $0.0393^{***}$ | $0.0350^{***}$ | $0.0078^{***}$ | $0.0141^{***}$ | $0.0222^{***}$ | $0.0245^{***}$ | | |
| | (23.53) | (25.39) | (6.91) | (14.64) | (15.14) | (14.63) | | |
| G | $0.0189^{***}$ | $0.0283^{***}$ | $0.0075^{***}$ | $0.0039^{***}$ | $0.0121^{***}$ | $0.0124^{***}$ | | |
| | (10.74) | (19.51) | (6.29) | (3.75) | (8.25) | (7.37) | | |
| Reference | Non-labeled | Non-labeled | Non-labeled | Non-labeled | Non-labeled | Non-labeled | | |
| Controls | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| County | X | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Municipality | X | X | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Zip code | X | X | X | ✓ | ✓ | ✓ | ✓ | ✓ |
| Unit FE | X | X | X | X | ✓ | ✓ | ✓ | ✓ |
| Observations | 747,387 | 747,387 | 747,387 | 747,006 | 177,634 | 118,627 | 177,634 | 92,416 |
| Adjusted $R^2$ | 0.478 | 0.644 | 0.762 | 0.819 | 0.968 | 0.957 | 0.968 | 0.970 |
| RMSE | 0.427 | 0.353 | 0.288 | 0.252 | 0.108 | 0.108 | 0.108 | 0.099 |

Note: Table 4 reports the cross-sectional OLS, HDFE and panel FE estimation results for the sales data. The dependent variable is the natural logarithm of the sales price. The default for EPC-labels is non-labeled dwellings. Control variables stand for hedonic characteristics (Table 2) and time FE (year, quarter, and month). Fixed location effects for county, municipality, and zip code. The sample of panel FE (column 5) includes 18 counties, 394 municipalities and 2587 zip codes (100%, 93%, and 53% of the whole of Norway, respectively). The sample period spans from January 2010 to December 2017. Heteroskedasticity robust *t*-statistics are in parentheses. $^{*}$ $p < .05$, $^{**}$ $p < .01$, $^{***}$ $p < .001$.

Next, we consider energy efficiency as a continues variable in columns 7 and 8, using the whole- and labeled-only sample respectively. The results show that an improvement in energy efficiency by one letter is associate with about 0.8-0.9% premiums on average, which is similar to the corresponding premium of energy efficiency improvement in the rental market. Further, we apply the same Heckman selection model as in the rental market analysis. The results of this robustness estimation, reported in the Appendix, Table A2, show that the coefficient of the selection variable is positively significant, indicating that unobserved factors behind the adoption of labeling is associated with higher sales prices. However, the coefficients of the energy efficiency premiums remain unchanged.

Table 6: Heterogeneity analyses of the impact of energy efficiency on sales prices

| | Largest cities | Largest cities excluded | Apartment | Detached | Other types | Small $\leq 60$ m$^2$ | Medium $> 60 \leq$ 120 m$^2$ | Large $> 120$ m$^2$ |
|---|---|---|---|---|---|---|---|---|
| Efficiency | 0.0078$^{***}$ | 0.0054$^{***}$ | 0.0081$^{***}$ | 0.0023$^{*}$ | 0.0041$^{***}$ | 0.0086$^{***}$ | 0.0057$^{***}$ | 0.0025$^{**}$ |
| | (15.07) | (9.26) | (18.70) | (2.46) | (3.99) | (13.05) | (11.55) | (2.90) |
| Controls | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Observations | 98,916 | 78,718 | 135,087 | 29,930 | 11,954 | 59,379 | 86,517 | 31,738 |
| Adjusted $R^2$ | 0.965 | 0.972 | 0.962 | 0.982 | 0.977 | 0.949 | 0.962 | 0.982 |
| RMSE | 0.104 | 0.109 | 0.112 | 0.085 | 0.076 | 0.120 | 0.099 | 0.079 |

Note: The dependent variable is the natural logarithm of the sales price. Controls stand for hedonic characteristics (Table 2), year, quarter, and month fixed effects (FE) and unit FE. The largest Norwegian cities are defined as having a population exceeding 90,000 (Oslo, Bergen, Stavanger, Sandnes, Trondheim, Fredrikstad, Sarpsborg, Drammen, and Porsgrunn). Heteroskedasticity robust *t*-statistics are in parentheses. $^{*}$ $p < .05$, $^{**}$ $p < .01$, $^{***}$ $p < .001$.

Further, we conduct the same heterogeneity analyses as for the rental market, reported in Table 6. Compared with the rental market, there is greater variation across the different stratifications in the sales market. We find that premiums in the largest cities are higher than for the rest of Norway, with 0.8% and 0.5%, respectively. Regarding dwelling types, there are lower premiums for detached homes and other types, compared with apartments. Interestingly, the relationship between premiums and dwelling size seems to be opposite of that in the rental market; here, larger dwellings have lower energy efficiency premiums. A reason for this could be that the energy savings in larger homes are relatively lower, as a percentage of total price, than for smaller dwellings. This is also corroborated by the lower premiums of detached homes, which in general are larger than apartments. Hence, since energy efficiency improvements are relatively less costly for smaller dwellings, there is less economic incentive for improvements in larger homes, which could help explain the energy efficiency paradox, although we do not observe this in the rental market.

Overall, the results indicate that all energy efficiency coefficients are subject to locational bias in the rental and sales markets. The dwellings with lower energy efficiency are associated with more locational bias in the rental market, while this bias is higher for the energy efficient dwellings in the sales market. Further, we find that the lower the energy efficiency, the less bias comes from unobserved quality in the sales market. All results are robust to potential bias originating from sample selection using the Heckman two-step model. The heterogeneity analysis also provides further evidence that different stratifications of the data may yield varying results. The findings therefore suggest that using nonrepresentative samples, not sufficiently controlling for unobserved locational heterogeneity, or conducting analyses of only a submarket may generate inferences that do not reflect the true price impact of energy efficiency for the whole market.

## 3.2 Label adoption impact on prices

So far, we have investigated the price effect of energy efficiency regarding the various endogeneity sources using both cross-section HDFE and panel FE to control for location heterogeneity and time-invariant factors. Nevertheless, there is still a potential endogeneity issue regarding time-variant factors in the error term that may be correlated with the labels. The IV/two-stage least squares (2SLS) method is the most common approach used in the literature to handle this endogeneity issue. However, it is challenging to apply the IV method using the specification of energy efficiency from the previous section. Instead, we therefore investigate the signaling effect of label adoption on prices, i.e., whether the information provided by the label itself has a price impact.

The idea behind the IV approach is that using a valid IV allows us to decompose the endogenous variable into a problematic component that may be correlated with the error term and another problem-free component (exogenous) that is uncorrelated with the error term that can be used in the structural model to obtain unbiased estimates. To accomplish this, we need at least one valid instrument for each endogenous variable that satisfies two conditions: the instrument relevance and the instrument exogeneity conditions.

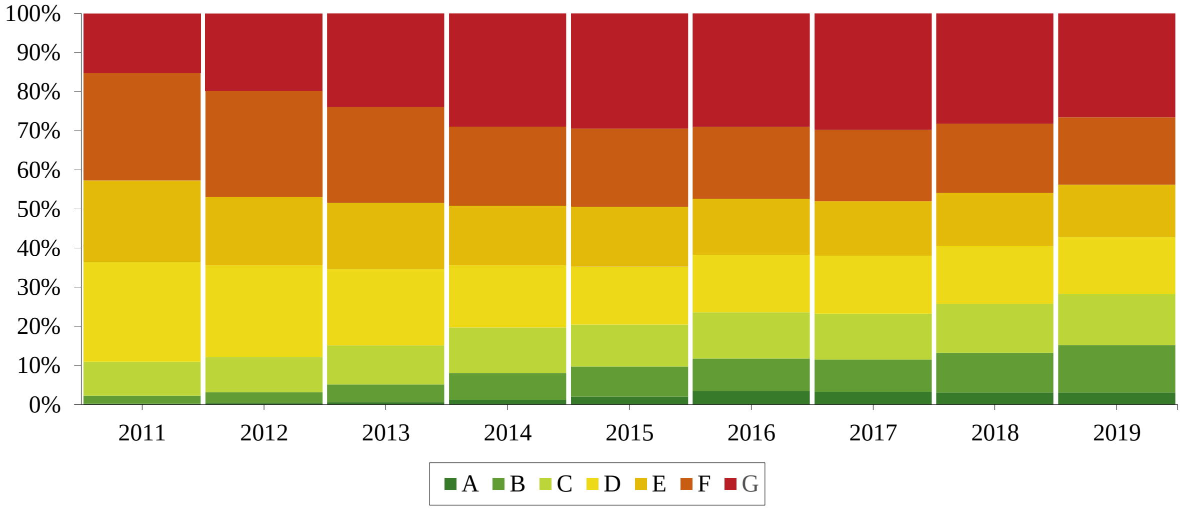


Figure 1: The distribution of labels among new energy performance certificates over time
Note: Distribution of labels among new energy performance certificates for existing Norwegian dwellings over time. For each year, the distribution of new certificates aggregates to 100%. Although dwellings built after 2010 should have at least a C-label, NNC also includes older dwellings that have not been labeled before that time and may therefore have lower energy efficiency. Data source: Energimerking.no.

We use the total monthly number of new EPCs (NNC) in absolute terms at the municipality level, obtained from Energimerking.no, as an instrument for the suspected endogenous variable *Label,* which is a dummy variable taking the value 1 if the dwelling is labeled, and 0 otherwise – the coefficient of this variable could be also considered a weighted average of the energy efficiency premium coefficients.[1] The instrument exogeneity condition requires independence from both the unobserved determinants of price and the measurement error (i.e., $Cov(NNC, u) = 0$). This assumption cannot be tested, as the coefficients in our model are exactly identified; however,

[1] Energimerking.no is the Norwegian online platform for EPC building registration managed by the Ministry of Climate and Environment.

because the NNC is randomly assigned regardless of the price level, there is no reason to believe that the number of new certificates can influence prices directly, but rather only through labels.[2] Figure 1 shows the distribution of new certificates over time and illustrates that new labels are not only comprised of new dwellings, which should have at least a C-rating, but also older dwellings that have not been labeled earlier. Although the EPC can be issued at any time, regardless of whether the owner intends to sell or rent out the dwelling, it could also be argued that the instrument may be picking up market liquidity – hence, violating the exogeneity condition – if a share of EPCs are effectively only issued for sales or rental transactions. We, therefore, include an interaction term between municipality and year and quarter as an additional fixed effect to mitigate this potential issue.[3] Additionally, this interaction term helps to control for any potential differences in submarket trends.

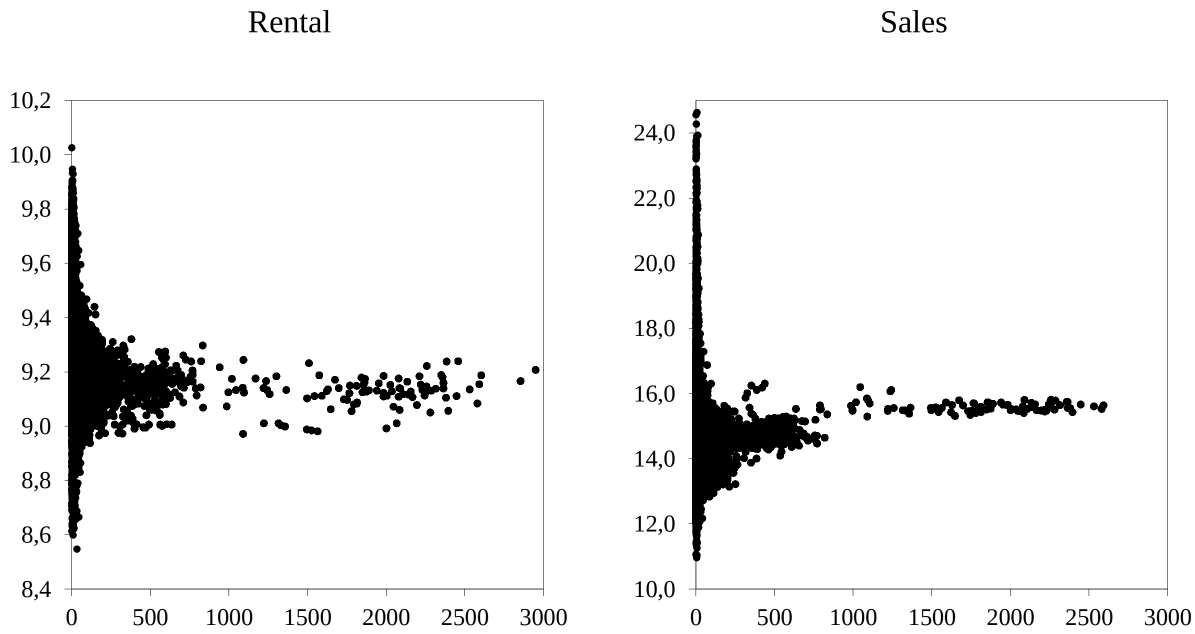


Figure 2: Instrument and price plots
Note: The monthly number of new certificates (NNC) against the monthly average log rental and sales prices.

[2] For multiple instruments, this assumption can be partially tested by applying the overidentification restriction test.
[3] We thank an anonymous referee for pointing out this potential issue.

The scatter plots in Figure 2 show that there is little to no relation between NNC and prices, which helps support the assumption of exogeneity. The instrument relevance condition requires that the IV must be correlated with the endogenous variable (i.e., $Cov(NNC, Label) \neq 0$). One way to test this assumption is to compute the *F*-statistic, testing the hypothesis that the coefficient on the instrument, $\beta$, is equal to zero in the first-stage regression, and the rule is that the *F*-statistic should exceed 10 to avoid weak instrument issues.[4]

Because the endogenous variable *Label* is a dummy variable, the conventional IV procedure cannot be applied directly, as *Label* is likely to be correlated with a binary conditional expectation function in the first stage of the conventional IV procedure. Thus, we adjusted our IV estimation following Angrist and Pischke (2008, pp. 142–144) by regressing *Label* on the instrument variable *NNC* and the covariates using a Poisson pseudo-maximum likelihood estimation with HDFE:

$$Label_{icmz} = exp\ (\sigma NCC_m + X_{icmz}\,\gamma\ + \alpha_c + \mu_m + \omega_z + \lambda_t + \epsilon_{icmz}), \tag{3}$$

where *Label* follows a Poisson distribution with equal mean and variance, and the remainder is as defined in equation (1). Next, we obtain the nonlinear fitted values, $\widehat{Label_{icmz}}$, from equation (3) to be added as an instrument in the conventional IV procedure.[5] The first and second-stage regression models of the IV estimation can be written as follows:

$$Label_{icmz} = \pi(\widehat{Label_{icmz}}) + \beta(NCC_m) + X_{icmz}\,\gamma\ + \alpha_c + \mu_m + \omega_z + \lambda_t + e_{icmz}, \tag{4}$$

$$ln\ price_{icmz} = \tau(\widehat{\widehat{Label_{icmz}}}) + X_{icmz}\,\gamma\ + \alpha_c + \mu_m + \omega_z + \lambda_t + u_{icmz}, \tag{5}$$

[4] A heteroskedasticity robust standard error should be used to obtain a reliable *F*-statistic when testing the relevance condition in the first-stage IV estimation.

[5] Angrist and Pischke (2008) applied a Probit regression, but we apply a Poisson pseudo-maximum likelihood estimation because it allows us to include multiple locational fixed effects.

where $\widehat{Label}_{icmz}$ is the fitted value of *Label* from the first-stage regression, also known as reduced-form regression, represented in equation (4), and $\tau$ is the IV regression consistent estimator of the true label effect on prices, conditional on the satisfaction of both IV assumptions.

### 3.2.1 Rental market results

Columns 1 and 2 in Table 7 present equation (1) estimations for the rental data by OLS and IV regression, respectively, without controlling for locational effects. We know that the label coefficient is biased upwards due to unobserved location effects, and this problem is not solved by applying IV regression. To reduce both sources of potential bias, unobserved locational and quality factors, we re-estimate equation (1) by applying HDFE and IV-HDFE in specifications 3 and 4, respectively. We found that the label coefficients are not significantly different, indicating that the main source of endogeneity is due to locational heterogeneity rather than unobserved quality. The results from Table 7 indicate that applying 2SLS without sufficient controls for locational effects does not solve the endogeneity issue (column 2). A violation of the second IV assumption is a potential explanation for this because it is likely that the instrument is correlated with the unobserved locational variation in the error term. The *F*-statistic of the IV-HDFE first-stage regression (column 4) is about 1,105, indicating that the instrument condition of relevance holds. We further apply the Durbin–Wu–Hausman test for endogeneity by plugging the residuals from the IV-HDFE first-stage regression into the HDFE model, and the residual coefficient *t*-statistic of 6.3 indicates that the IV-HDFE estimation is preferred.

Table 7: Identifying the unobserved location and quality heterogeneity for rental prices

| | OLS | IV | HDFE | IV-HDFE |
|---|---|---|---|---|
| Label | 0.0766*** | 8.2406*** | 0.0421*** | 0.0410*** |
| | (81.28) | (36.21) | (65.84) | (60.35) |
| *Control variables* | ✓ | ✓ | ✓ | ✓ |
| *Fixed location effects* | X | X | ✓ | ✓ |
| *Excluded instrument* | X | *NNC* | X | *NNC* |
| Observations | 669,894 | 669,366 | 669,448 | 665,002 |
| Adjusted $R^2$ | 0.403 | 0.531 | 0.725 | 0.724 |
| RMSE | 0.286 | 0.253 | 0.194 | 0.194 |
| First stage $F$-statistic | | 1,147.18 | | 1,104.90 |

Note: Table 7 reports the ordinary least squares (OLS), high-dimensional fixed effects (HDFE), instrumental variable (IV) estimation results for the rental data. The dependent variable is the natural logarithm of the monthly rental price. The default is non-labeled dwellings. Control variables stand for hedonic characteristics (Table 1) and time FE (year, quarter, and month). Fixed location effects for county, municipality, and zip code. The IV, number of monthly new EPCs (*NNC*), is at the municipality level. The IV first-stage $F$-statistic is robust for heteroskedasticity. The sample period spans from January 2011 to September 2019. Heteroskedasticity robust $t$-statistics are in parentheses. $^{*}$ $p < .05$, $^{**}$ $p < .01$, $^{***}$ $p < .001$.

### 3.2.2 Sales market results

Specifications 1 and 2 in Table 8 report the OLS and IV estimations of the signaling effect of label adoption on sales prices. The OLS result indicates that, holding other factors constant, the labeling effect at any rating is associated with a 6.1% sales price premium without considering locational FE. Like the rental estimations (Table 7), applying IV estimation without locational controls does not provide reasonable estimates for the sales data. Applying the HDFE in specification 3, the signaling effect of label adoption has a premium of 3%. Comparing the $R^2$ of the OLS and HDFE specifications demonstrates that the locational FE explain about 34% of the variation in prices. Applying the IV-HDFE, the label coefficient does not change significantly compared with the HDFE specification. Finally, the two-panel FE estimations with and without IV provide further evidence of the signaling effect of label adoption on prices with premiums of 2.1% to 2.3%.

The panel IV-FE is expected to produce the most consistent estimates because the FE feature manages time-invariant heterogeneity and the IV feature manages other sources of heterogeneity.

Nevertheless, the result from the panel IV-FE is not substantially different from the IV-HDFE estimation. The excluded instrument *F*-statistic of 132 from the reduced form regression indicates that the relevance condition is satisfied, and the Durbin–Wu–Hausman endogeneity test with a *t*-statistic of 4.24 provides evidence in favor of the panel IV-FE regression.

Table 8: Identifying the unobserved location and quality heterogeneity for sales prices

| | OLS | IV | HDFE | IV-HDFE | Panel FE | Panel IV-FE |
|---|---|---|---|---|---|---|
| Label | 0.0612*** | 6.657*** | 0.0303*** | 0.0297*** | 0.0228*** | 0.0208*** |
| | (51.12) | (58.56) | (40.87) | (35.12) | (20.07) | (17.19) |
| *Control* | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| *Fixed location effects* | X | X | ✓ | ✓ | ✓ | ✓ |
| *Panel fixed effects* | X | X | X | X | ✓ | ✓ |
| *Excluded instrument* | X | *NNC* | X | *NNC* | X | *NNC* |
| Observations | 747,387 | 745,774 | 747,006 | 745,247 | 177,634 | 177,422 |
| Adjusted $R^2$ | 0.474 | 0.525 | 0.817 | 0.795 | 0.968 | 0.968 |
| RMSE | 0.429 | 0.401 | 0.253 | 0.259 | 0.108 | 0.109 |
| Excluded instrument *F*-statistic | | 1,772.41 | | 691.69 | | 131.79 |

Note: Table 8 reports the ordinary least squares (OLS), high-dimensional fixed effects (HDFE), instrumental variable (IV), and panel FE estimation results for the sales data. The dependent variable is the natural logarithm of the sales price. The default is non-labeled dwellings. Control variables stand for hedonic characteristics (Table 2) and time FE (year, quarter, and month). Fixed location effects for county, municipality, and zip code. The IV, number of monthly new EPCs (*NNC*), is at the municipality level. The IV first-stage *F*-statistic is robust for heteroskedasticity. The sample period spans from January 2010 to December 2017. Heteroskedasticity robust *t*-statistics are in parentheses. * $p < .05$, ** $p < .01$, *** $p < .001$.

So far, we have used various identification strategies to establish the signaling effect of label adoption on market prices and rents, addressing the importance of locational heterogeneity and unobserved quality. The overall findings suggest that the source of endogeneity depends on the market and the methods used in the analyses. The label premiums from the rental market are mainly subject to bias originating from locational heterogeneity, while in the sales market both locational and unobserved quality are potential sources. Using the IV method without sufficient control for location may not help provide reliable result. The findings also show that controlling for the lowest unit, panel fixed effect, is preferable to reveal the true signaling effect of labeling adoption, however, alternatively using the IV method with sufficient control for location may be a potential solution.

### 3.2.3 Label adoption impact on prices over time

Next, we investigate the signaling effect of label adoption over time in terms of immediate, short-run, and long-run effects by recursively estimating the IV-HDFE specifications of equation (5) for the rental and sales data. This section aims to illustrate that the time window may produce different results, which consequently may yield different conclusions and policy implications. We first examine Norway as a whole, then only the largest cities, and finally, the rest of Norway separately, using a one-month window starting from January 2010 and January 2011 for the sales and rental data, respectively.

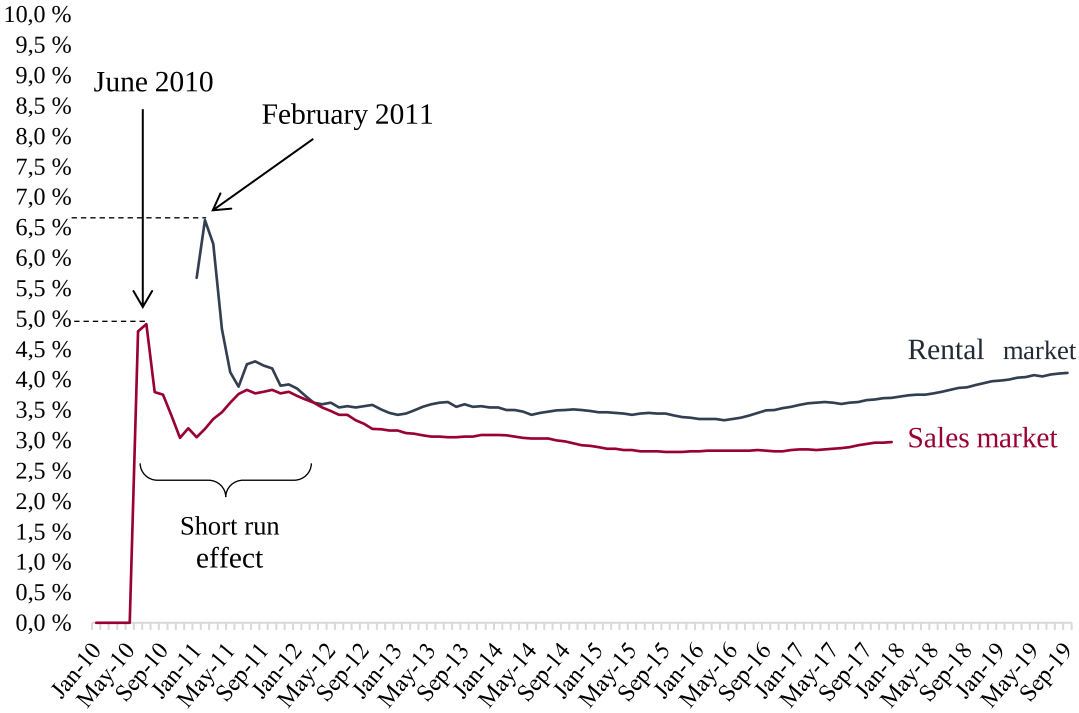


Figure 3: Recursive instrumental variable (IV) estimations over time, the whole of Norway

Note: Figure 3 reports the recursive IV-estimated label coefficients (equation 5) for the Norwegian sales and rental markets over time. We use a one-month window starting from January 2010 for sales data (96 estimations) and from January 2011 for rental data (105 estimations). The EPC implementation date in Norway was July 1st, 2010

Figure 3 reports the recursively estimated label coefficients for the whole of Norway over time. Although 224 labeled dwellings ranged from A to G in our sample before the implementation date in Norway of July 1, 2010, no significant effect in the sales market was found until June, with a peak in July. This indicates that the signaling effect of label adoption had a positive immediate effect at the time of effectuation. In the rental market, the peak was observed in February 2011 and is relatively higher than in the sales market. The respective peaks are followed by a decline and a period of instability, representing the short-run effects. The figure reveals that the effects occurs earlier in the sales market than the rental market. In the long run, the label effect on both markets seems to stabilize, and the rental effect is consistently higher.

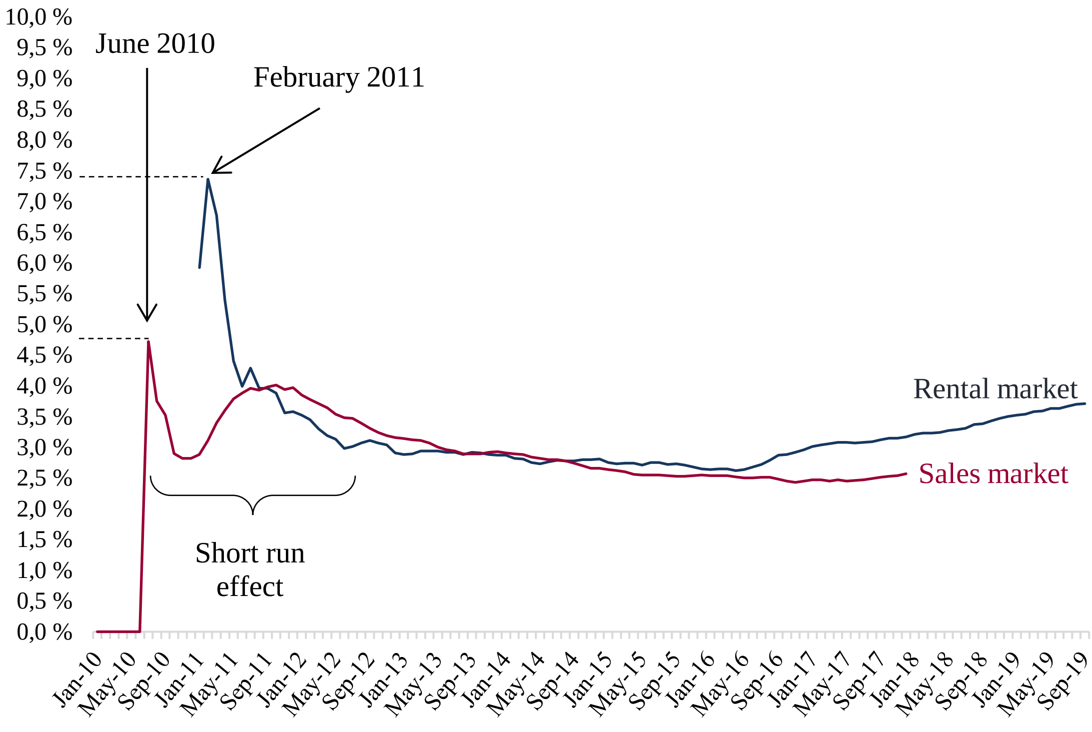


Figure 4: Recursive instrumental variable (IV) estimations over time, largest cities

Note: Figure 4 reports the recursive IV-estimated label coefficients (equation 5) for the largest Norwegian sales and rental markets over time. The largest cities are defined as having a population exceeding 90,000 inhabitants (Oslo, Bergen, Stavanger, Sandnes, Trondheim, Fredrikstad, Sarpsborg, Drammen, and Porsgrunn). We use a one-month window starting from January 2010 for sales data (96 estimations) and from January 2011 for rental data (105 estimations). The EPC implementation date in Norway was July 1st, 2010.

In Figure 4, we report the recursively estimated equation (5) for the largest cities, defined as having a population exceeding 90,000. Compared with the whole of Norway, the rental market peak is higher, whereas the sales market peak is approximately the same. The short-run effects are similar, but a stronger decline is observed in the rental market. In the long run, the effects on both markets are similar in magnitude but are lower overall compared with the whole of Norway. Next, we exclude the largest cities and investigate the rest of Norway in Figure 5. The market peaks are relatively equal, and the effects seem to stabilize at a faster rate compared with the largest cities. Additionally, the rental market effect is consistently higher than that of the sales market.

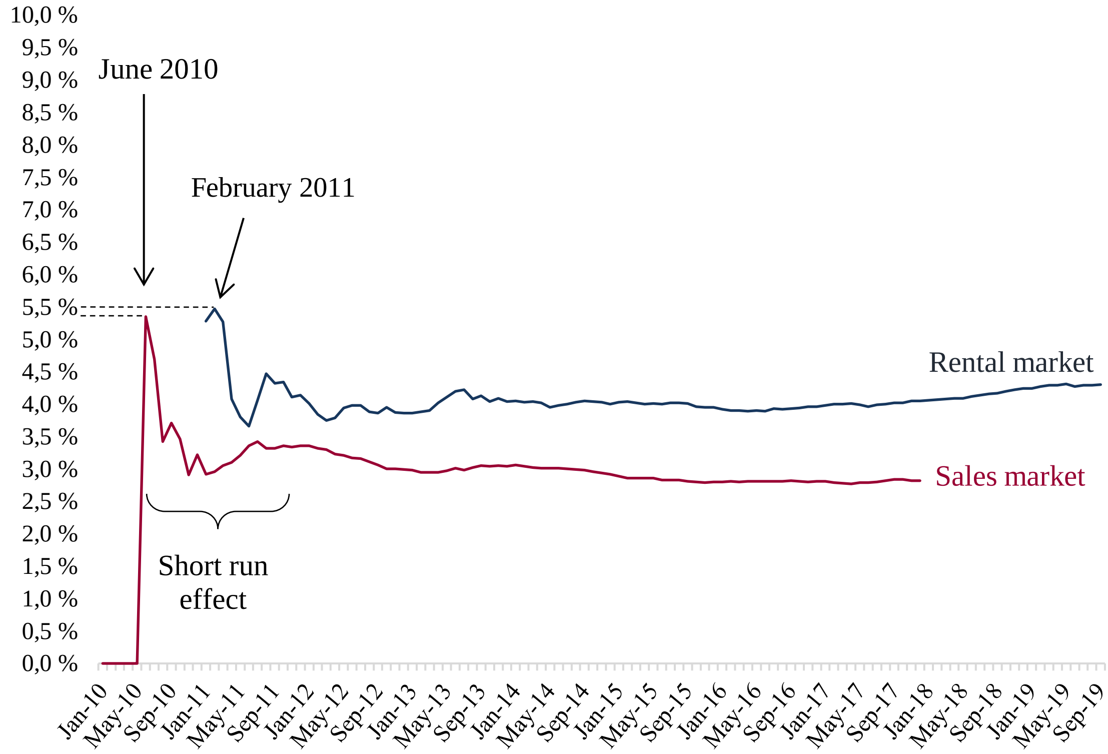


Figure 5: Recursive instrumental variable (IV) estimations over time, rest of Norway
Note: Figure 5 reports the recursive IV-estimated label coefficients for sales and rental markets (equation 5) of Norway excluding the largest Norwegian cities, defined as having a population exceeding 90,000 inhabitants (Oslo, Bergen, Stavanger, Sandnes, Trondheim, Fredrikstad, Sarpsborg, Drammen, and Porsgrunn). We use a one-month window starting from January 2010 for sales data (96 estimations) and from January 2011 for rental data (105 estimations). The EPC implementation date in Norway was July 1st, 2010.

Overall, the recursive estimations of the signaling effect of label adoption indicate that an immediate positive effect occurred in the sales market, followed by short-run instability before stabilizing in the long-run. In the rental market, we observe the same pattern, but with a slower response and a higher overall influence. Further, more short-run variation was found in the large-city sample compared with the more stable effect observed in the rest of Norway. These findings highlight the possibility that different conclusions might be drawn due to sample selection issues related to location, time periods and submarkets.

# 4 Conclusion

Through the implementation of energy performance certificates for buildings in 2002, and more recent endeavors such as the CONSEED and PENNY projects, funded by Horizon 2020, the EU has a continuing focus on energy consumption, efficiency and development of policies to increase environmental sustainability. The capitalization of energy efficiency in transaction prices and rents has been subject to much research in recent years. However, it is challenging to assess the policy implication of EPC implementation due to several confounding factors, such as issues related to data limitation and sample time period, together with endogeneity problems originating from omitted variables that might be correlated with energy efficiency and prices.

This paper addresses the issues related to the price impact of energy efficiency by applying different identification strategies using highly representative samples from the Norwegian rental market ($N$ = 670,000) between 2011 and 2019 and the Norwegian sales market ($N$ = 750,000) between 2010 and 2017. First, we thoroughly control for locational effects by employing both the HDFE approach for the pooled cross-sectional data and panel FE to identify the source of endogeny in terms of unobserved locational effects. Second, we apply the IV approach using the total number of new labeled dwellings as an instrument to assess the signaling effect of label adoption on prices. Finally, we investigate the labeling effect over time by recursively estimating the IV regression over the respective sample periods.

We provide evidence that the valuation of energy efficiency is subject to endogeneity originating from unobserved locational factors, and that dwellings with lower energy efficiency are associated with more locational bias in the rental market, while this bias is higher for the energy efficient dwellings in the sales market. Further, we find that the lower the energy efficiency, the less bias comes from unobserved quality in the sales market. Overall, improving the energy efficiency of the dwelling with one letter on the EPC rating has similar effects for both rental and sales objects, with a price impact of about 0.8-1.0%.

The premiums of energy efficiency in our results support earlier findings from the Norwegian rental market (Khazal and Sønstebø, 2020), and are also very much in line with existing studies from the rest of Europe. Generally, the price impact of energy efficiency on rental dwellings seems to be of the same magnitude throughout. In the sales market analysis, however, we reach different conclusions about energy efficiency premiums compared with earlier Norwegian studies (Olaussen et al., 2017, Olaussen et al., 2019), but our findings support the majority of existing studies from the EU. In these studies, the premiums associated with energy efficiency are generally similar to ours.

Investigating the signaling effect of label adoption, we discover that the IV approach produces biased estimations if the location is not controlled for. However, after considering this issue, both the IV and panel FE demonstrate that a positive signaling effect exists on prices in both markets. In the rental market, the IV approach provides evidence that the main source of endogeneity is due to locational heterogeneity rather than unobserved quality standards. The recursive estimations reveal that labeling has immediate, short-run, and long-run price effects and that different effects are observed in different submarkets. The findings also highlight the possibility that different conclusions might be drawn due to sample selection issues related to time periods and submarkets, and that methodological and data limitations are essential factors that must be considered when assessing the effects of the EPC implementation.

# 5 References


ANGRIST, J. D. & PISCHKE, J.-S. 2008. *Mostly harmless econometrics: An empiricist's companion,* Princeton university press.

AYDIN, E., BROUNEN, D. & KOK, N. 2017. Information asymmetry and energy efficiency: Evidence from the housing market. *Technical report, Maastricht University Working Paper*.

BROUNEN, D. & KOK, N. 2011. On the economics of energy labels in the housing market. *Journal of Environmental Economics and Management,* 62**,** 166-179.

BRUEGGE, C., CARRIÓN-FLORES, C. & POPE, J. C. 2016. Does the housing market value energy efficient homes? Evidence from the energy star program. *Regional Science and Urban Economics,* 57**,** 63-76.

CAJIAS, M., FUERST, F. & BIENERT, S. 2019. Tearing down the information barrier: the price impacts of energy efficiency ratings for buildings in the German rental market. *Energy Research & Social Science,* 47**,** 177-191.

CAJIAS, M. & PIAZOLO, D. 2013. Green performs better: energy efficiency and financial return on buildings. *Journal of Corporate Real Estate,* 15**,** 53-72.

CERIN, P., HASSEL, L. G. & SEMENOVA, N. 2014. Energy performance and housing prices. *Sustainable Development,* 22**,** 404-419.

CHEGUT, A., EICHHOLTZ, P. & HOLTERMANS, R. 2016. Energy efficiency and economic value in affordable housing. *Energy Policy,* 97**,** 39-49.

CHEGUT, A., EICHHOLTZ, P., HOLTERMANS, R. & PALACIOS, J. 2019. Energy Efficiency Information and Valuation Practices in Rental Housing. *The Journal of Real Estate Finance and Economics***,** 1-24.

COURT, A. 1939. *Hedonic price indexes with automotive examples, in "The dynamics of automobile demand"*, General Motors, New York, pp. 98-119.

DAVIS, P. T., MCCORD, J. A., MCCORD, M. & HARAN, M. 2015. Modelling the effect of energy performance certificate rating on property value in the Belfast housing market. *International Journal of Housing Markets and Analysis*.

DE AYALA, A., GALARRAGA, I. & SPADARO, J. V. 2016. The price of energy efficiency in the Spanish housing market. *Energy Policy,* 94**,** 16-24.

DRESSLER, L. & CORNAGO, E. 2017. The rent impact of disclosing energy performance certificates: Energy efficiency and information effects. *ECARES Working Papers*.

EICHHOLTZ, P., KOK, N. & QUIGLEY, J. M. 2010. Doing well by doing good? Green office buildings. *The American Economic Review,* 100**,** 2492-2509.

FEIGE, A., MCALLISTER, P. & WALLBAUM, H. 2013. Rental price and sustainability ratings: which sustainability criteria are really paying back? *Construction Management and Economics,* 31**,** 322-334.

FREGONARA, E., ROLANDO, D. & SEMERARO, P. 2017. Energy performance certificates in the Turin real estate market. *Journal of European Real Estate Research*.

FUERST, F. & MCALLISTER, P. 2011a. Green noise or green value? Measuring the effects of environmental certification on office values. *Real Estate Economics,* 39**,** 45-69.

FUERST, F. & MCALLISTER, P. 2011b. The impact of Energy Performance Certificates on the rental and capital values of commercial property assets. *Energy Policy,* 39**,** 6608-6614.

FUERST, F., MCALLISTER, P., NANDA, A. & WYATT, P. 2015. Does energy efficiency matter to home-buyers? An investigation of EPC ratings and transaction prices in England. *Energy Economics,* 48**,** 145-156.

FUERST, F., MCALLISTER, P., NANDA, A. & WYATT, P. 2016a. Energy performance ratings and house prices in Wales: An empirical study. *Energy Policy,* 92**,** 20-33.

FUERST, F., OIKARINEN, E. & HARJUNEN, O. 2016b. Green signalling effects in the market for energy-efficient residential buildings. *Applied Energy,* 180**,** 560-571.
GAURE, S. 2013. OLS with multiple high dimensional category variables. *Computational Statistics & Data Analysis,* 66**,** 8-18.
GUIMARAES, P. & PORTUGAL, P. 2010. A simple feasible procedure to fit models with high-dimensional fixed effects. *The Stata Journal,* 10**,** 628-649.
HAN, L. & STRANGE, W. C. 2016. What is the role of the asking price for a house? *Journal of Urban Economics,* 93**,** 115-130.
HECKMAN, J. J. 1979. Sample selection bias as a specification error. *Econometrica: Journal of the econometric society***,** 153-161.
HYLAND, M., ALBERINI, A. & LYONS, S. 2016. The Effect of Energy Efficiency Labeling: Bunching and Prices in the Irish Residential Property Market. Trinity College Dublin, Department of Economics.
HYLAND, M., LYONS, R. C. & LYONS, S. 2013. The value of domestic building energy efficiency—evidence from Ireland. *Energy Economics,* 40**,** 943-952.
HÖGBERG, L. 2013. The impact of energy performance on single-family home selling prices in Sweden. *Journal of European Real Estate Research*.
HÅRSMAN, B., DAGHBASHYAN, Z. & CHAUDHARY, P. 2016. On the quality and impact of residential energy performance certificates. *Energy and buildings,* 133**,** 711-723.
JENSEN, O. M., HANSEN, A. R. & KRAGH, J. 2016. Market response to the public display of energy performance rating at property sales. *Energy Policy,* 93**,** 229-235.
KHAZAL, A. & SØNSTEBØ, O. J. 2020. Valuation of energy performance certificates in the rental market–Professionals vs. nonprofessionals. *Energy Policy,* 147**,** 111830.
KHOLODILIN, K. A., MENSE, A. & MICHELSEN, C. 2017. The market value of energy efficiency in buildings and the mode of tenure. *Urban Studies,* 54**,** 3218-3238.
KOK, N. & JENNEN, M. 2012. The impact of energy labels and accessibility on office rents. *Energy Policy,* 46**,** 489-497.
LYONS, R. C. 2019. Can list prices accurately capture housing price trends? Insights from extreme markets conditions. *Finance Research Letters,* 30**,** 228-232.
MUDGAL, S., LYONS, L., COHEN, F., LYONS, R. & FEDRIGO-FAZIO, D. 2013. Energy performance certificates in buildings and their impact on transaction prices and rents in selected EU countries. *European Commission (DG Energy): Paris, France*.
MURPHY, L. 2014. The influence of the energy performance certificate: The Dutch case. *Energy Policy,* 67**,** 664-672.
OLAUSSEN, J. O., OUST, A. & SOLSTAD, J. T. 2017. Energy performance certificates–Informing the informed or the indifferent? *Energy Policy,* 111**,** 246-254.
OLAUSSEN, J. O., OUST, A., SOLSTAD, J. T. & KRISTIANSEN, L. 2019. Energy Performance Certificates—The Role of the Energy Price. *Energies,* 12**,** 3563.
RAMOS, A., PÉREZ-ALONSO, A. & SILVA, S. 2015. Valuing energy performance certificates in the Portuguese residential sector. *Economics for Energy. WP***,** 02-2015.
ROSEN, S. 1974. Hedonic prices and implicit markets: product differentiation in pure competition. *Journal of Political Economy,* 82**,** 34-55.
SALVI, M., HOREHAJOVA, A. & NESSER, J. 2010. *Der Nachhaltigkeit von Immobilien einen finanziellen Wert geben: Der Minergie_Boom unter der Lupe. Eine Marktanalyse der ZKB, CCRS, Universität Zürich,* Zürich, Dr. Erika Meins (Hrsg).
SKOGSTAD, H. P. 2011. Leieprisstatistikk for leiligheter i Oslo - metode og praktisk gjennomføring, i "Gjengs leie dokumentasjonsrapport". (Rental price statistics for apartments in Oslo - methods and practical implementation, in "Common rents documentation report"). Boligbygg Oslo KF.

STANLEY, S., LYONS, R. C. & LYONS, S. 2016. The price effect of building energy ratings in the Dublin residential market. *Energy Efficiency,* 9**,** 875-885.

TALTAVULL, P., ANGHEL, I. & CIORA, C. 2017. Impact of energy performance on transaction prices. *Journal of European Real Estate Research*.

VON GRAEVENITZ, K. & PANDURO, T. E. 2015. An alternative to the standard spatial econometric approaches in hedonic house price models. *Land Economics,* 91**,** 386-409.

WAHLSTRÖM, M. H. 2016. Doing good but not that well? A dilemma for energy conserving homeowners. *Energy Economics,* 60**,** 197-205.

WILEY, J. A., BENEFIELD, J. D. & JOHNSON, K. H. 2010. Green design and the market for commercial office space. *The Journal of Real Estate Finance and Economics,* 41**,** 228-243.

WILHELMSSON, M. 2019. Energy performance certificates and its capitalization in housing values in Sweden. *Sustainability,* 11**,** 6101.

ZHANG, L., LI, Y., STEPHENSON, R. & ASHURI, B. 2018. Valuation of energy efficient certificates in buildings. *Energy and Buildings,* 158**,** 1226-1240.

# 6 Appendix

Table A1: Summary of literature

| | Co. | Sample | Period | Methods | Major findings |
|---|---|---|---|---|---|
| Khazal and Sønstebø (2020) | NOR | 440,000 (R) | 2011-2018 | Multilevel and Heckman approaches includes county, municipality and zip-code fixed effects | Labeled dwellings have a premium, and the premium is increasing with a higher EPC-label |
| Olaussen et al. (2019) | NOR | 4,693 (S) | 2000-2014 | OLS and panel fixed effects include dummies for city districts (Oslo) | Energy label and energy performance of dwellings has little to no effect on transaction prices. |
| Wilhelmsson (2019) | SWE | 100,000 (S) | 2013-2018 | Propensity score methods, quantile regression with county and municipality effects | EPCs are not differently capitalized in the high-end housing price segment |
| Cajias et al. (2019) | GER | 1,029,202 (R) | 2013-2017 | Generalized Additive Model for Location, Scale and Shape with regional dummies | Small but significant rental premium for green dwellings. |
| Chegut et al. (2019) | ENG<br>NLD | 12,000 (R)<br>53,000 (R) | 2012, 2015<br>2010, 2015 | OLS includes postcodes fixed effects (North West England and Amsterdam) | In England, no premium for energy performance 2012, but a premium exists in 2015. In the Netherlands, premium in 2010, and larger premium in 2015. |
| Zhang et al. (2018) | USA | 12,000 (S) | 2007-2010 | OLS includes dummies for census (Atlanta) | Homes with energy certificates were sold with a premium of almost 12% |
| Olaussen et al. (2017) | NOR | 2,066 (S) | 2009-2014 | OLS and panel fixed effects include dummies for city districts (Oslo) | Premium for energy efficiency, not for labeling. |
| Aydin et al. (2017) | NLD | 30,036 (S) | 2008-2011 | OLS, IV and repeat-sales methods with neighborhood dummies and location of the home relative to city center | Energy efficiency is capitalized into prices, but not the labeling itself. |
| Fregonara et al. (2017) | ITA | 879 (S) | 2011-2014 | OLS includes geographical segmentation characteristic of cities | No label impact on prices. |
| Kholodilin et al. (2017) | GER | 7,298 (S)<br>13,366 (R) | 2011-2014 | OLS and Heckman approaches include locational characteristics and distance to city center (Berlin) | Energy savings are generally capitalized in prices and rents. |
| Dressler and Cornago (2017) | BEL | 6,262 (R) | 2010-2014 | OLS, 2SLS and Heckman approaches include municipality dummies (Brussels, 19) | Highly energy efficient dwellings achieve a premium of 4.8% compared to inefficient dwellings. |
| Taltavull et al. (2017) | ROU | 16,420 (S) | 2013-2015 | IV and GLS approaches include dummies for city districts (Bucharest) | Premiums of 2.2-6.5% for retrofitted green dwellings. |
| Wahlström (2016) | SWE | 69,698 (S) | 2009-2010 | OLS and includes yearly number of frost days per municipality | No price premium for energy efficient housing. |
| Hårsman et al. (2016) | SWE | 75,726 (S) | 2009-2010 | OLS and includes Neighborhood characteristics & distance to shoreline | Selling prices do not increase by a higher EPC efficiency rating. |
| Fuerst et al. (2016a) | WAL | 191,544 (S) | 2003-2014 | OLS and repeat sales regressions include postcode fixed effects and urban/rural indicator | Premium of 12.8% for A/B-label compared with D-label. |
| Fuerst et al. (2016b) | FIN | 6,194 (S) | 2009-2012 | OLS includes postcode fixed effects and CBD distance | Price premium for ABC-labeled dwellings. |

Table A1 (cont.): Summary of literature

| | Co. | Sample | Period | Methods | Major findings |
|---|---|---|---|---|---|
| Chegut et al. (2016) | NLD | 17,835 (S) | 2008-2013 | OLS includes postcodes fixed effects | Highly energy efficient dwellings sell for 2.0-6.3% more than similar dwellings with low energy efficiency. |
| Jensen et al. (2016) | DEN | 117,483 (S) | 2007-2012 | OLS with postcode fixed effects | Energy performance ratings of buildings have an impact on property sales prices |
| Stanley et al. (2016) | IRL | 2,792 (S) | 2009-2014 | OLS includes postcodes fixed effects (Dublin, 8 regions) | Energy efficiency has a positive effect on residential property list prices. |
| Bruegge et al. (2016) | USA | 6,000 (S) | 1998-2009 | OLS includes 50 small subdivision-level indicator variables (Florida) | Price premium for new labeled homes, but the effect is smaller after a few years |
| Fuerst et al. (2015) | ENG | 333,095 (S) | 1955-2012 | OLS and repeat sales regression include postcode fixed effects and dummies of 8 urban levels | A/B-rated dwellings achieve a premium of 5% compared to D-rated dwellings. |
| Davis et al. (2015) | IRL | 3,797 (S) | | OLS (Belfast) no location controls | A small but positive relationship between better energy performance and higher selling prices. |
| Ramos et al. (2015) | PRT | 21,170 (S) | 2007-2014 2015 | OLS and Heckman approaches include location attributes and dummies for district | Portuguese consumers have a high valuation for high rated dwellings. |
| Cerin et al. (2014) | SWE | 64,753 (S) | 2009-2010 | OLS includes dummy for urban metropolitan areas | Energy performance associated with a price premium. |
| Högberg (2013) | SWE | 1,073 (S) | 2009 | OLS includes neighborhood specific attributes (Stockholm) | Price premium for energy efficiency. |
| Hyland et al. (2013) | IRL | 15,060 (S) 20,825 (R) | 2008-2011 | OLS and Heckman selection model with location characteristics | Price premium of 9.3% (sales) and 1.8% (rentals) for A-label, and 5.5% (sales) and 3.9% (rentals) for B-label compared with D-label. |
| Cajias and Piazolo (2013) | GER | 2,615 (S, R) | 2008-2011 | OLS with location dummies, latitude and longitude | Premium of 0.45% (sales) and 0.08% (rentals) per 1% increase in energy savings. |
| Feige et al. (2013) | CHE | 2,453 (R) | 2009-2011 | OLS includes 7 dummies for location with different quality (location 1=best location) | Positive association between sustainable features and rental prices. |
| Kok and Jennen (2012) | NLD | 1,100 (R) | 2005-2010 | OLS including ZIP code fixed effects | A premium of 6.5 percent associated with energy efficient buildings |
| Brounen and Kok (2011) | NLD | 177,000 (S) | 2008-2009 | Probit and Heckman two-step includes province-fixed effects and neighborhood characteristics | Premium of 10% for A-label and 5.5% for B-label compared with D-label. |
| Fuerst and McAllister (2011a) | USA | 9,806 (S) 18,519 (R) | 1999-2008 2008 | OLS with location (longitude & latitude) | Labeled buildings have both a rental and sale price premium, compared to buildings in the same submarkets. |

Note: Co. indicates country of study; sample indicates the number of observations and whether the study investigates the sales market (S) or the rental market (R).

Table A2: Heckman selection models for rental and sales prices

| | Rental | | Sales | |
|---|---|---|---|---|
| | Labeled-only sample | Heckman model | Labeled-only sample | Heckman model |
| A | 0.0470*** | 0.0465*** | 0.0762*** | 0.0724*** |
| | (22.25) | (21.85) | (3.56) | (3.36) |
| B | 0.0390*** | 0.0380*** | 0.0313*** | 0.0292*** |
| | (18.13) | (17.58) | (3.79) | (3.54) |
| C | 0.0356*** | 0.0353*** | 0.0374*** | 0.0363*** |
| | (17.70) | (17.46) | (7.95) | (7.72) |
| D | 0.0233*** | 0.0235*** | 0.0275*** | 0.0263*** |
| | (11.85) | (11.87) | (8.34) | (8.06) |
| E | -0.0008 | -0.0012 | 0.0147*** | 0.0139*** |
| | (-0.38) | (-0.57) | (6.24) | (5.97) |
| F | -0.0094*** | -0.0093*** | 0.0080*** | 0.0075*** |
| | (-4.27) | (-4.20) | (4.49) | (4.26) |
| G | | | | |
| Mills ratio | | -0.0248** | | 0.4139*** |
| | | (-2.69) | | (19.80) |
| Observations | 108,276 | 106,715 | 92,416 | 92,338 |
| Adjusted $R^2$ | 0.760 | 0.759 | 0.970 | 0.970 |
| RMSE | 0.165 | 0.165 | 0.099 | 0.099 |

Note: Table A2 reports the cross-sectional HDFE estimation for the rental data and panel FE estimation for the sales data. The dependent variable is the natural logarithm of the monthly rental price or the sales price. The default for EPC-labels is G-labeled dwellings. Heteroskedasticity robust *t*-statistics are in parentheses. * $p < .05$, ** $p < .01$, *** $p < .001$.