Using Estimates of Housing Energy Efficiency from Administrative Microdata to Measure Underconsumption and Energy Poverty

Lydia Geijtenbeek[1], Peter Mulder[23], Robin Niessink[2]

ABSTRACT

Accurately measuring the energy efficiency of housing is essential for determining the benefits of residential energy retrofitting, evaluating the effectiveness of energy efficiency policies, and assessing the extent and causes of energy poverty. Many EU countries use Energy Performance Certificates (EPCs) – or energy labels – to assess housing energy efficiency. However, this data is often outdated or incomplete. Moreover, even up-to-date EPCs are only weakly correlated with actual energy consumption, a discrepancy known as the performance gap. These issues make EPCs less well suited for evaluating housing quality or energy poverty.

In the context of measuring energy poverty, we address these issues by designing and implementing a model that estimates housing energy efficiency by combining EPC data with additional administrative resources. Our model builds upon related research that uses Dutch microdata for scenario studies or energy poverty, improving on this by using additional data sources and a finely tuned model. Our resulting model explains 51% of the variation in energy expenditure based on housing characteristics, as opposed to 40% for EPCs. We illustrate the application of our model in assessing energy poverty, in the form of the LILEQ indicator (i.e. Low Income Low Energy Quality), and demonstrate that this indicator correlates more closely with survey-based consensual indicators of energy poverty (e.g. EU-SILC), than common indicators that are based on actual expenditure (e.g. share of energy expenditure). Finally, we demonstrate additional uses for our model estimates, such as the identification of energy underconsumption or monitoring of housing quality.

In summary, our contribution lies in a calibrated model of energy consumption, which improves upon EPCs, and in demonstrating useful applications for this model, especially as input for a high quality measure of energy poverty.

Keywords

Housing Energy Efficiency, Energy Performance Certificates (EPCs), Energy Labels, Energy Poverty, Microdata, Energy policy evaluation

Acknowledgements

We like to thank Manon Middelkoop, Ruud van den Wijngaart, Anika Batenburg, Charlotte Brand, Francesco Dalla Longa, Krista Keller, Reinder Lok, Lindsey van der Meer and Koen Straver from CBS and TNO for their respective contributions to building a framework and comprehensive dataset on measuring energy poverty in The Netherlands over several years.

[1] Statistics Netherlands (CBS), corresponding author: l.geijtenbeek@cbs.nl.
[2] Netherlands Organization of Applied Scientific Research (TNO).
[3] Utrecht University, Copernicus Institute of Sustainable Development.

1. Introduction

Background and literature

In response to ambitious climate targets and persistently high energy prices, both governments and individual citizens across the European Union (EU) have intensified efforts to improve energy efficiency within the residential sector. As part of this strategy, the revised Energy Performance of Buildings Directive (EU/2024/1275) mandates a 16% improvement in the energy performance of residential buildings by 2030. This requirement reflects the dual importance of energy efficiency: reducing greenhouse gas emissions and mitigating the socio-economic consequences of high energy costs. Residential energy use significantly contributes to both global emissions and household expenditure, thereby affecting income distribution, particularly in the context of rising energy prices. Consequently, energy efficiency has become a focal point for policies at European, national, and regional levels. Accurate, granular, and timely data on energy efficiency, energy use, and energy poverty are essential for designing and evaluating such policies.

Energy performance certificates (EPCs) or energy labels are widely used to assess housing energy efficiency. However, in practice, these labels often suffer from incompleteness and inaccuracy. For instance, in the Netherlands, approximately 39% of dwellings lacked a valid energy label in 2024 (CLO 2025), while existing EPCs are not always updated following renovations, and are thus often outdated or inaccurate (Van den Wijngaart and van Polen 2020). Furthermore, a significant body of empirical research has demonstrated that EPCs often do not reflect actual residential energy use. This mismatch, known as the energy performance gap, undermines the utility of EPCs for policymaking, household decision-making, and energy poverty measurement.

Majcen (2016) provided an early but relevant study using Dutch microdata, finding that theoretical EPC-based energy use projections consistently exceeded actual household consumption, particularly in low-efficiency dwellings. Although dated, her work remains foundational in the Dutch context. Additional studies reinforce these conclusions. Melis et al. (2023) assessed EPC reliability across housing stock in Ireland and found that EPCs explain only a modest portion of the variance in measured heat loss. Few et al. (2023), in a comprehensive study of the EPC-consumption gap in Great Britain using data from 1,300 homes with both energy label information and actual consumption records, found that EPCs systematically overpredict energy use for homes with lower EPC bands with the largest discrepancies in the lowest-rated homes. Notably, the performance gap persisted even in homes with behavioral conditions aligned with EPC model assumptions—such as standard occupancy and thermostat use—suggesting that technical or structural assumptions, rather than behavior alone, drive much of the mismatch. This finding is echoed in Van Hove et al. (2022), who analyzed over 100,000 homes in Flanders and found substantial deviations between modeled and actual energy use. Their study also emphasizes multicollinearity and model overspecification as important contributors to EPC inaccuracies. While the former studies relate to the static performance gap, there is also a dynamic performance gap in respect with retrofitting. In this context, Peñasco and Díaz Anadón (2023), Allcott and Greenstone (2017), and Fowlie et al. (2018) highlight the systematic overestimation of energy savings following retrofitting measures based on EPC-derived predictions.

In response to the limitations of EPCs, researchers have turned to data-driven models that estimate actual energy consumption using housing and household characteristics. These approaches—ranging from classical statistical models to advanced machine learning algorithms—have demonstrated superior predictive performance and practical scalability. Amasyali and El-Gohary (2018) provide a widely cited review of building-level energy prediction models. They distinguish between physics-based ‘forward models’ that simulate energy use using detailed physical and environmental data, and data-driven models that rely on historical consumption patterns and statistical learning. While the former require extensive, often inaccessible inputs, data-driven models can leverage administrative microdata to make reliable predictions across large housing stocks. Recent advances in machine learning further support this shift. Cui et al. (2024) find that models incorporating structural characteristics and household composition yield high predictive accuracy across residential building types.

Similarly, Harputlugil and De Wilde (2021) emphasize that human-building interaction is essential for understanding residential energy consumption, which models using occupant and socio-economic data can incorporate more easily than EPCs. These models often use features such as dwelling type, floor area, insulation presence, heating system, construction year, and renewable energy installations (e.g., solar panels). For example, research by Van den Wijngaart and Van Polen (2020), followed up by van Beijnum et al. (2023) use a regression model based on publicly available dwelling features to predict regional residential heat demand. Their approach served as an inspiration to our approach.

Energy poverty, broadly defined as the inability to attain adequate energy services at affordable cost, is traditionally measured by expenditure thresholds. Common instances are a share of energy expenditure above 10% or twice the median, as used by the European Commission (Rademaekers et al. 2016). Yet these indicators receive criticism for not excluding households with high incomes and strong preferences for energy services above socially accepted standards (Al Kez et al. 2024). Furthermore, they do not account for housing efficiency and therefore fail to distinguish between low energy consumption due to either effective insulation or financial hardship (Mulder et al. 2023). Several alternative frameworks aim to integrate housing quality in energy poverty measurements. The UK's Low Income Low Energy Efficiency (LILEE) measure, for example, combines income with EPC bands A–C as a threshold (Department for Business & Energy and Industrial Strategy, 2022). Yet, as Semple et al. (2024) show, this approach still has limitations: in low-income neighborhoods in London, 27% of households in homes rated EPC A–C reported experiencing energy insecurity, indicating underestimation of energy poverty. Karpinska and Smiech (2023) go a step further by constructing empirical typologies of energy-poor households using actual energy use data. They argue that reliable poverty assessment requires linking income, housing costs, and energy use, thereby emphasizing the need for robust energy efficiency estimation; something EPCs fail to provide consistently.

Goal and scope

This paper addresses the limitations discussed above of both EPCs and energy poverty indicators, by proposing a model that aims to estimate housing energy efficiency more accurately than EPCs, and to apply this to the assessment of energy poverty and underconsumption.

For this model we integrate EPCs with a broad range of administrative data, thus building upon previous models that use Dutch microdata to estimate housing quality above or in the absence of EPCs (e.g. Van den Wijngaart and Van Polen 2020, Van Beijnum et al. 2023, Mulder et al. 2023). Similarly to these models, we aim to generate robust estimates of energy efficiency that remain unaffected by occupant behavior, while ensuring that the model can be applied to most of the Dutch housing stock. To achieve this, we use data available consistently across all Dutch homes from 2019 at Statistics Netherlands (CBS).

Our model differs in a few respects from similar models. First, since our main context is energy poverty, our dependent variable is total net energy expenditure, and not energy consumption or energy use for space heating. This means that our measure of housing quality does not only depend on insulation levels, but also on energy generation (e.g. by solar panels), and the share of electricity in energy consumption. Second, although we base our estimates primarily on characteristics of a dwelling, we also control for social factors of the inhabitants. Third, we use additional variables such as the presence of solar panels and main heat source. And finally, we use a sparser and more finely tuned specification with fewer degrees of freedom.

The model presented in this paper builds on ongoing collaboration between CBS and TNO to monitor and analyze energy poverty in the Netherlands. Central to this work is the development of the LEQ (Low Energy efficiency Quality) indicator, designed to identify households living in poorly insulated dwellings—a key structural driver of energy poverty. Introduced as part of a broader monitoring framework established in 2020, LEQ supports more accurate diagnosis and targeting of vulnerable households. This framework has informed national policy through an annual regionally detailed assessment of energy poverty across the country and is detailed in a series of publications, including Statistics Netherlands (2025), Batenburg et al. (2024) and Mulder et al. (2023).

The remainder of this paper is structured as follows: Section 2 describes the data sources and model design; Section 3 demonstrates applications to energy efficiency and poverty analysis; Section 4 discusses limitations and implications; while Section 5 concludes.

2. Methodology

Data and variable selection

This research uses the rich administrative data set that is available to researchers at Statistics Netherlands (CBS). The development and estimation of our model was based on a population of 7,8 million houses in 2019, for which a large number of attributes is available.

The dependent variable in all of our estimations is the total *energy expenditure* at an address. Energy expenditure is calculated by using the net supply of gas and electricity to any home in the Netherlands, and combining this with average gas and electricity prices. The resulting energy expenditure includes both fixed and variable costs of energy, taxes, revenues from electricity generated by solar panels, and estimates for district heating. Note that our estimates are thus not restricted to heating costs (as in some studies), but include all of the energy costs and revenues attributable to a certain home.

Out of the vast available data at Statistics Netherlands, we considered both physical characteristics of dwellings (e.g., floor area, construction period) and socio-economic characteristics of households (e.g., disposable income, household size) that were found to be related to energy consumption in other research, or were otherwise deemed relevant. From other studies (e.g. Van Beijnum et al. 2023) we have learned that it is important to keep the number of distinct variables low, so that relevant interactions can be included without adding too many degrees of freedom to the model. Thus, we restricted the number of variables from an initial list of fifteen to eight by using regression tree analysis.

Table 1 shows the initial variable list and the outcomes of two regression trees. The first tree has energy expenditure as dependent variable (tree 1). We use the *importance* statistic (Ishwaran 2007) as our main selection criterium, and find that floor area has the largest importance. Yet, other variables with high importance are known to be related to floor area (e.g. property value), with a risk of including redundant or collinear regressors in our model. To prevent this, we did two things. First we measured how often a variable actually appears in the tree, by summing the number of nodes in the tree that depend on a certain variable, weighted by the fraction of cases in the node (see 'Sum of nodes' in Table 1). Second, we did a second regression tree analysis in which the dependent is energy expenditure corrected for area (tree 2). A variable with large importance in tree 1, yet a small node count and small importance in tree 2 is now deemed potentially collinear and less relevant. Our final variable selection consists of floor area, based on its large importance in tree 1, and People, Solar, BuildType, BuildYear, Label, HeatSource, and Income based on their importance and weighted sum of nodes in tree 2 (both with an arbitrary threshold of 0,3). Among the variables that are excluded are PropertyValue, due to its assumed collinearity with area, and type of ownership, which was nevertheless included in the model by Van Beijnum et al. (2023) and Mulder et al. (2023).

Table 1 Variables that were considered for the model, with indicators of relevance derived from on two different regression trees. Of these, tree 1 has energy expenditure as dependent variable, and tree 2 has energy expenditure corrected for floor area as dependent variable. For both trees we report the 'Importance' statistic, together with a weighted 'sum of nodes' of a certain variable in the tree. The *final column shows which variables were selected as regressors in the model. Bold face numbers indicate that they were used as positive selection criterium.*

		Tree 1: expenditure		Tree 2: expenditure, corrected for area		
Variable name	Description of variable	Importance	Sum of nodes	Importance	Sum of nodes	Selected
Area	Floor area of the house (in m^2)	**1,00**	1,57	0,14	**0,41**	**yes**
People	Number of inhabitants	0,35	1,19	**0,82**	**1,46**	**yes**
Solar	Presence of solar panels (yes/no)	0,23	0,66	**1,00**	**1,00**	**yes**
BuildType	One of: detached house, semi-detached house, corner house, terraced house, apartment	0,79	1,25	**0,50**	**0,99**	**yes**
BuildYear	Year of construction of the dwelling	0,14	0,39	**0,57**	**1,17**	**yes**
Label	Certified energy label (EPC): A (incl., A+, A++, ...), B, C, D, E, F, or No label.	0,10	0,27	0,27	**0,36**	**yes**
HeatSource	Main heat source. One of: individual gas heating, block heating, district heating, individual electric heating	0,13	0,18	0,23	**0,33**	**yes**
Income	Disposable household income (at address), in percentiles.	0,38	0,00	**0,44**	0,00	**yes**
PropertyValue	Taxable property value of the house	0,49	0,00	0,21	0,11	no
HasGas	House had a connection to the gas grid; either for heating or (also) warm water, and/or cooking	0,04	0,00	0,09	0,00	no
InhabitYears	Number of years that the current residents live in the house	0,04	0,00	0,05	0,00	no
Ownership	Ownership type of the house	0,13	0,00	0,03	0,00	no
LabelYear	Year that the label (EPC) was issued	0,10	0,00	0,02	0,00	no
IncomeSource	Main income source of the household (e.g. wage, profits, benefits for unemployment, social security, or sickness, pension, etc.)	0,02	0,00	0,02	0,00	no
Age	Age of main resident of the household	0,03	0,00	0,02	0,00	no

Model building

For the selected variables we evaluated the shape of the relationship with energy expenditure, and found strong non-linearities for most numerical variables. Therefore, most variables are included in the model as categorical variables, with categories that mirror the observed non-linearities. Only area appears roughly linear, yet with an upper bound at about 300m^2. Since area is such a strong predictor, we assessed other specifications than purely linear, such as loglinear, polynomial and piece wise linear, and found that a higher degree polynomial or piece-wise linear function best captures the slight S-shape of the relation between area and energy expenditure. In the final model we use a piece-wise linear form for area combined with capped linear form to be used in interactions with other variables. Table 2 lists the variables that were eventually included in the model, along with a description of their selected form.

Table 2 Variables that were selected for our regression model for predicting energy expenditure, along with a description of their functional form in the regression. Readers that want to know the exact shape (including cut-offs) will receive this information on request by the first author.

Variable name	Specification and form of the variable
Area	Linear, capped from above at 302m^2; for use in interactions with other variables.
AreaPWL	Piece wise linear (PWL), in 12 pieces with selection of split points based on significance in a separate regression model
BuildPeriod	Categorical, in 9 periods, determined by a separate regression model, which reflect changes in historical building regulations.
BuildType	Categories: detached house, semi-detached house, corner house, terraced house, apartment, unknown
Label	Categories: A (incl., A+, A++, ...), B, C, D, E, F, or No label
Solar	Boolean (true/false)
HeatSource	Categories: individual gas heating, block heating, district heating, individual electric heating, unknown
People	Categories 0, 1, 2, 3, 4, and 5 or more
Income	Categorical, in 9 income brackets, based on percentiles of disposable household income. The relation is roughly U-shaped

Both the regression tree and existing research indicate that interactions of variables are potentially relevant for the prediction of energy consumption. Van den Wijngaart and Van Polen (2020) use a full 4-way interaction of all their variables (i.e. building period, EPC label, house type, floor area), while Van Beijnum et al. (2023) use a 5-way interaction of their set of variables that adds type of ownership, while Mulder et al (2023) include a 3-way interaction of building period, house type and floor area. This resulted in large models that are computationally hard to estimate, and are potentially prone to overspecification.

For the current model we only considered two-way interactions. We evaluated several instances, ranging from no interactions at all to including all two-way interactions, on data from 2020. The main results are summarized in Table 3. The final model (see also Equation 1) attempts to strike a balance between accuracy and sparsity.

Table 3 Comparison of different versions of the model, which differ in the number of two-way interactions that were included.

Description of the model	Degrees of freedom	Model quality (adjusted R2)
No interactions, only separate variables	50	0,566
All interactions between two variables	649	0,586
Selected set of interactions, see equation 1	137	0,579

The resulting model includes both main effects and selected two-way interactions that were found to significantly improve explanatory power. The full specification of the model is given in Equation 1 below.

Equation 1 Formal specification of our model, in which energy costs are a function of characteristics of both the house and its inhabitants.

$$EnergyCost = \alpha + \beta_1\, Area + \beta_2\, AreaPWL + \beta_3\, Solar + \beta_4\, BuildPeriod + \beta_5 Label + \beta_6\, HeatSource + \beta_7\, BuildType + \beta_8\, Income + \beta_9\, People + \beta_{10}\, Area * Solar + \beta_{11}\, Area * Income + \beta_{12}\, Area * HeatSource + \beta_{13}\, Area * People + \beta_{14}\, Area * Label + \beta_{15}\, Area * BuildPeriod + \beta_{16}\, Solar * BuildType + \beta_{17}\, Solar * HeatSource + \beta_{18}\, Solar * People + \beta_{19}\, Solar * Label + \beta_{20}\, HeatSource * People + \beta_{21}\, BuildYear * Label$$

Our final model can explain more than half of the variation in energy expenditure (R^2=0.584) in our base year 2019, and is thus reasonably effective. Table 4 shows the combined significance of sets variables and interactions. Due to the substantial number of coefficients (i.e., 138), we do not include a full table of estimated parameters. We find that, although parameter estimates for single categories are sometimes not significant ($p > 0.10$), most of the combined effects of the dummy variables that together define a certain categorical variable or interaction are all highly significant ($p < 0.01$).

Since the p-values are all very low, we use F-scores to determine which variables are most influential in the regression. Here exists a clear trade-off between having a large F-score and being included in many interactions, with more interactions leading to a lower F-score. For instance, building type has the largest F-score while only being included in a single interaction, while area is included in six interactions and has the lowest F-score. Among the most influential predictors are the type of dwelling, number of residents, presence of solar panels in interaction with floor area, and construction period. Notably, interaction effects between solar presence and housing characteristics (e.g., area, label, build type) suggest non-uniform impacts of energy improvements across the housing stock, possibly due to differences in electrical capacity of solar panels that can be installed on certain building types.

Table 4 Combined significance, F-score and degrees of freedom for variables (e.g. Solar), sets of variables that make up a categorical variable (e.g. BuildType is a set that consists of five dummy variables for each of the categories detached house, semi-detached house, corner house, terraced house, apartment, with unknown being the reference category), or interactions.

Variable, set of variables, or interaction between two (sets of) variables	Degrees of freedom	F-score	Combined significance
Intercept	1	4306,561	0,000
Solar	1	4624,407	0,000
Area (linear, incl. dummy for unknown)	2	3,799	0,024
AreaPWL (piece wise linear)	12	70,925	0,000
BuildPeriod	8	4234,216	0,000
Label	7	856,989	0,000
BuildType	5	101862,046	0,000
Income	8	561,676	0,000
People	4	9115,203	0,000
HeatSource	4	858,785	0,000
Solar * Area	1	6442,281	0,000
Income * Area	8	3197,782	0,000
HeatSource * Area	4	546,129	0,000
People * Area	4	1507,040	0,000
Label * Area	5	481,251	0,000
BuildPeriod * Area	5	2374,237	0,000
BuildType * Solar	5	3456,916	0,000
HeatSource * Solar	4	1348,715	0,000
People * Solar	4	1240,410	0,000
Label * Solar	5	1448,728	0,000
People * HeatSource	16	550,522	0,000
BuildPeriod * Label	25	921,223	0,000
Total regression	138	66866,826	0,000

Applications

In its raw form, this model can be used to estimate annual energy expenditure of a certain combination of characteristics for a house and its inhabitants. This is relevant for predictions or scenario studies. Yet in this paper we use the model for two other types of estimates.

First, we abstract from household characteristics and floor area to obtain an estimate for the energy efficiency of a dwelling, which we define as the energy expenditure that is inherent to a building. To isolate this intrinsic property from occupant behavior and household characteristics, we simulate each dwelling's energy cost using standardized values for floor area (100m^2), and number of residents (2 people), and income (median). This allows us to generate an adjusted efficiency score that reflects only the structural and technical attributes of the home. This efficiency score can serve as a robust alternative to EPC labels, which are often outdated or missing for a large share of the housing stock.

Second, we abstract from income to get a benchmark energy expenditure to be used for under- and overconsumption analysis. Here we compare actual and expected consumption. To this end we use our model to estimate energy expenditure based on both dwelling and household characteristics, yet abstracted from income, by replacing actual household income by the median income level. This is necessary because income and energy consumption are correlated, and including the income effect would set the bar too low for low-income households.

Thus, these results lay the empirical foundation for the applications discussed in Section 3, including the estimation of housing energy efficiency and the identification of energy poverty and behavioral underconsumption.

3. Results

Our model has several policy-relevant applications, which we explore in this section: estimating energy efficiency, identifying energy poverty, and detecting over- and underconsumption. Together, these three applications illustrate the versatility and policy relevance of the model, providing new insights into both structural inefficiencies and behavioral adaptations within the residential energy landscape.

Estimating energy efficiency

In this context, energy efficiency relates to the energy costs that are inherent to a building. As mentioned in the previous section, we calculate energy efficiency of a house by applying our model to the characteristics of the house, while replacing its actual floor area by the default of 100m^2, and supplying default values for the characteristics of its inhabitants (i.e. 2 persons with median income). This procedure results in an estimated yearly energy cost for the house, which is independent of its inhabitants or floor area. We will refer to this value as Energy Efficiency Score (or EEScore). Since this measure is continuous, there may be an unfair advantage with respect to the discrete EPC labels. To account for this and for fair comparison, we also created a discrete version of the EEScore in 8 levels.

In Table 5 we compare our model-based efficiency score with EPC labels. To this end, we model the actual energy expenditure as a linear combination of efficiency and floor area, and measure the predictive power of the model in the form of its R^2. We evaluated three models: one for energy labels (EPCs) in 8 categories, including a category 'No label', a second which uses our continuous EEScore, and a third which uses the discrete version of our measure. Table 5 demonstrates that our efficiency score explains a greater proportion of variation in energy use ($R^2 = 0.515$) than labels ($R^2 = 0.403$). This is not merely due to its continuous nature, since the discrete version of this variable with the same number of categories, still has a larger predictive power ($R^2 = 0.480$).

In other words: our modelled version of housing energy efficiency (ESScore) explains 51% of the variance in actual energy cost, versus 40% for using labels and size alone.

Table 5 Comparing prediction of energy consumption by labels versus estimated efficiency.

Energy efficiency indicator	Model specification	Degrees of freedom	R^2
EPC label, in 8 categories (A-G, and 'No label')	EnergyCost = Label + Label * Area	15	0,403
Estimated energy efficiency score	EnergyCost = EEScore + EEScore * Area	2	0,515
Estimated energy efficiency score, in eight categories of equal share	EnergyCost = EEScore_cat + EEScore_cat * Area	15	0,480

Table 6 indicates how EPC labels align with our model-based efficiency score, or energy quality. To this end we use two binary indicators that were derived from the efficiency score: Low Energy Quality (LEQ) coincides with the 50 percent least efficient homes, and Very Low Energy Quality (VLEQ) with the 15 percent least efficient homes, both with respect to the Dutch housing population in 2019. Table 6 shows how generally EPC labels and model-based efficiency quality are well-aligned, with labels A-C being most prevalent in non-LEQ homes, D-F labels in LEQ homes, and G labels in VLEQ homes. Also, while 53% of all homes do not have a valid EPC label in 2019, among the homes with the largest predicted energy costs (VLEQ) a majority of 87% does not have a valid EPC label. This means that identifying low energy efficient homes based on labels alone would exclude the majority of such homes. Also, the measures do not always align. For instance, of the homes classified by the model as inefficient, 1% have an EPC rating of A. This underlines the value of our method in identifying misclassifications in the existing energy labeling system.

Table 6. Relation between our estimate of energy efficiency and energy labels (EPC) in 2019.

			Share of houses in the row with a certain EPC energy label (%)							
Category of model-based efficiency score	Total	Share of total (%)	A	B	C	D	E	F	G	No label
Total housing stock	7.814.910	100	8	8	14	8	5	3	2	53
Non-Low Energy Quality (non-LEQ)	3.905.890	50	15	13	22	8	3	1	1	36
Low Energy Quality (LEQ)	3.909.030	50	1	2	6	9	6	4	3	69
Very Low Energy Quality (VLEQ)	1.196.680	15	0	1	2	2	2	3	4	87

Monitoring energy efficiency

The results of our model demonstrate its potential for policy-relevant monitoring of housing energy efficiency and energy poverty. One key application is tracking the evolution of energy quality in the housing stock over time.

Figure 1 illustrates the share of dwellings classified as low energy quality (LEQ) or very low energy quality (VLEQ) in municipalities in the Netherlands in 2019 and 2023. The figure shows a discernible trend to better quality; whereas in 2019 municipality averages were mostly between 50-60% for LEQ and 30-40% for VLEQ, in 2023 the most common ranges are 20-30% for LEQ and 10-20% for VLEQ. Due to the high prevalence of houses without a label, EPC labels cannot monitor the extend of this increase.

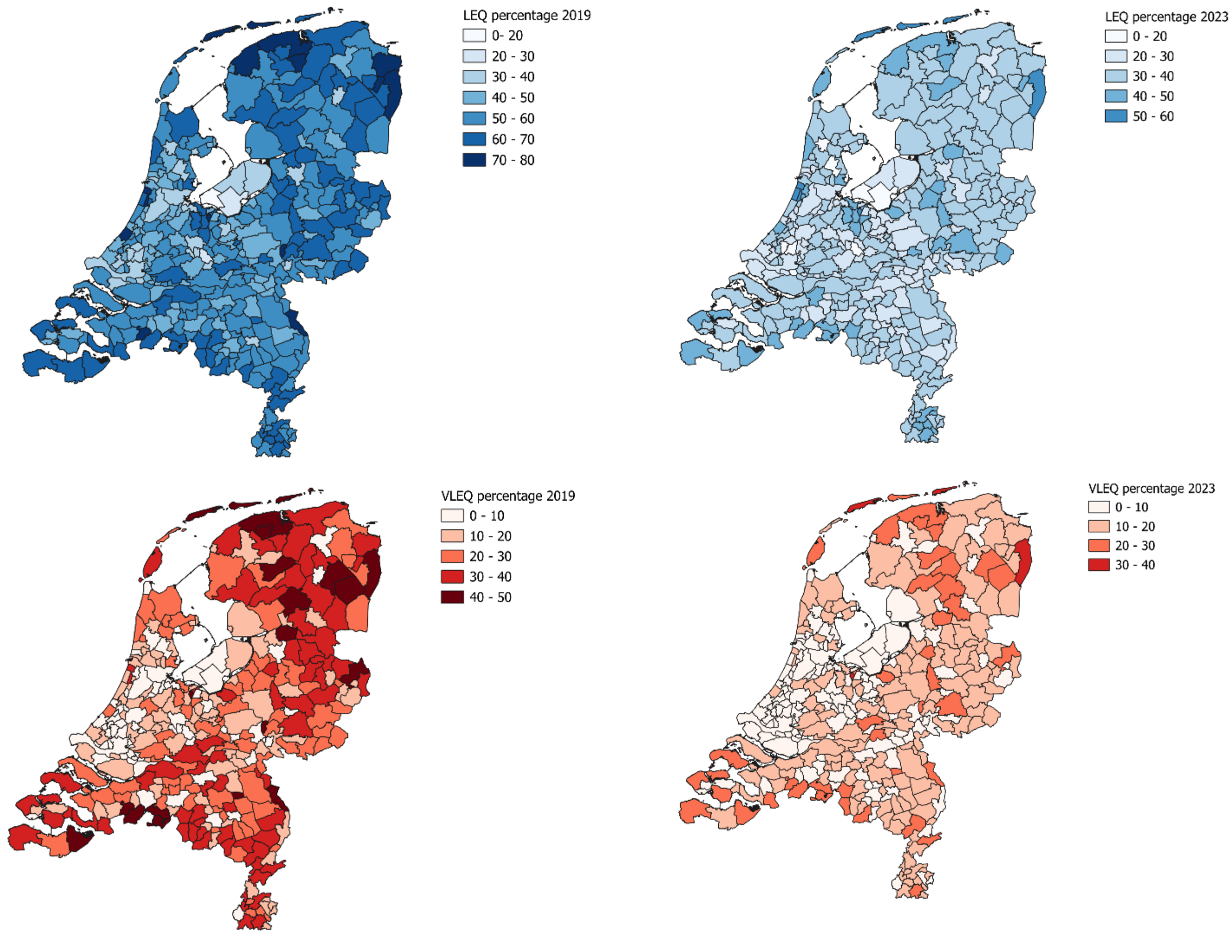

Figure 1 Share of (very) low energy quality, or (V)LEQ, houses in the Netherlands, 2019-2023. The top two panels show the share of low quality houses in 2019 (left) and 2023 (right), while the bottom two panels show the share of very poor quality houses. The share of non-LEQ houses is not shown separately, as it is the mirror image of LEQ.

An important caveat in using LEQ for assessing changes in housing efficiency is that not all increases can be detected by this measure. Only homes that have a different EPC label, newly installed solar panels, or a change in main heat source can be identified as having improved efficiency, apart from a small number of homes with a different LEQ-status due to administrative changes in floor area or construction year. Among the 1.4 million homes that were LEQ in 2019 and not in 2023, we saw the following changes: 74% had newly installed solar panels, 49% had a different label (including changes between an actual label and the category 'No label'), and 12% had a different main heat source, e.g. district heating, or electric heating.

Estimating energy poverty

Most expenditure base energy poverty indicators rely on observed energy expenditure in relation to income. However, this approach risks overlooking households who under-consume energy due to financial constraints, as low energy use can be misinterpreted as energy efficiency. By interpreting our model-based estimates as 'required energy expenditure', we can instead identify households whose structural energy needs are high but whose income is insufficient to meet those needs, irrespective whether they use much energy or not. This enables more accurate targeting of households experiencing "hidden" or structural energy poverty.

Table 7 presents a typical set of energy poverty indicators, including two indicators that use our Low Energy Quality (LEQ), as described in the previous section. TNO and CBS have developed a national framework for monitoring energy poverty, introduced in 2020 and updated annually (Statistics Netherlands 2025, Batenburg et al. 2024). This framework comprises multiple indicators, including HEQ (High Energy Quote), LIHE (Low Income

High Expenditure), and the LEQ-based indicator LILEQ (Low Income Low Energy Quality). Here follows a brief explanation:

- HEQ is calculated as having energy expenditure that is larger than 10% of household income. Since this would include many households with large wealth yet low income (such as entrepreneur-retirees), this indicator excludes households with wealth above the 90% percentile. This indicator is closely related to frequently used indicators based on the share of energy expenditure, such as the 10% or 2M indicators (Al Kez et al. 2024). Disadvantages were discussed in section 1.
- LIHE is defined as having an income below 130% of the poverty line, while having energy costs that are larger than the average cost of a label C dwelling in 2019. This measure is closely related to Low Income High Cost (LIHC).
- LILEQ is defined as having an income below 130% of the poverty line, while living in a house that has an energy efficiency among the 50% worst in 2019. Note that LEQ was specifically created to reflect housing quality as a structural factor in energy poverty. The LILEQ indicator is closely related to Low Income Low Energy Efficiency (LILEE) used in the UK. Both LILEE and LIHC were coined by the Department for Business & Energy and Industrial Strategy (2022).
- LIHELEQ is de combination of the previous two, including all households that have an income below 130% and either high energy costs (LIHE) or an inefficient home, or both. Statistics Netherlands and TNO use LIHELEQ as the main indicator for 'energy poverty'.

As outlined by Mulder et al. (2023), this monitoring system enables policymakers to identify energy-poor households according to different criteria, tuned for different policy perspectives. For instance, HEQ targets households that are sensitive to increases of energy prices, LIHE detects households that are at risk of falling below the poverty line when energy prices rise, while the use of LILEQ allows detection of households with low incomes who live in poorly insulated homes. Note that the monitoring excludes some households, with the largest groups being households with (partially) unknown income, student households, and households that share a common address. This leaves out roughly 1 million Dutch households (Statistics Netherlands, 2024).

Table 7 Overview of Dutch energy poverty indicators and the share of households that are considered energy poor according to each indicator, 2019.

Dutch energy poverty indicators	Count	Percentage
Total households for which energy poverty was measured	6.963.900	100
Households with high energy quote (HEQ)	513.500	7,4
Households with low income and high energy expenditure (LIHE)	436.800	6,3
Households with low income and low energy quality (LILEQ)	392.100	5,6
Households with low income and high energy expenditure *or* low energy quality (LIHELEQ)	601.700	8,6

Next to expenditure based indicators on energy poverty, policy and research also use consensual indicators, based on survey data (Al Kez et al, 2024). A commonly used set of such indicators is included in the European Union Statistics on Income and Living conditions (EU-SILC, 2019). These consist of i) Living in drafty, damp, or mold-prone homes; ii) Difficulty in maintaining an adequately warm indoor environment during winter; iii) Arrears on energy or utility bills. These indicators may serve as external validation points to test which administrative measures best identify households experiencing actual energy hardship.

Figure 2 shows how these consensual indicators relate to the three main types of expenditure-based indicators of Table 7: LEQ, LIHE and LILEQ. We use three different sources for these indicators. Survey data from the Dutch housing survey (Woon, 2018), and the European Union Statistics on Income and Living conditions (EU-SILC, 2019), complemented by administrative data on general arrears. The figure shows the percentage of positive answers to the questions as a percentage of all responses, aggregated at the address level.

We see that for these indicators the share of positive answers in the general population (left bar of each group) is relatively low, ranging from 1% with arrears on utility bills, to 25% with a drafty home. Yet, for each of the groups with energy poverty, the responses are markedly higher. This indicates that indeed groups that are labeled as energy poor bases on administrative data are more likely to actually undergo energy hardship. Furthermore, the results show that these effects are highest for LILEQ, for 6 out of 7 variables, with the exception of 'Total arrears', which is most common among energy poor according to LIHE. Which is indeed as expected, for people with low income and high (energy) expenditure are most likely to have problems with paying their bills. Meanwhile, households with HEQ, though commonly used, are much less likely to suffer from energy poverty related problems as mentioned in the survey-based indicators, which raises questions about its diagnostic value. These findings highlight the advantage of using model-based structural indicators like LEQ/LILEQ to capture the multidimensional nature of energy poverty more accurately.

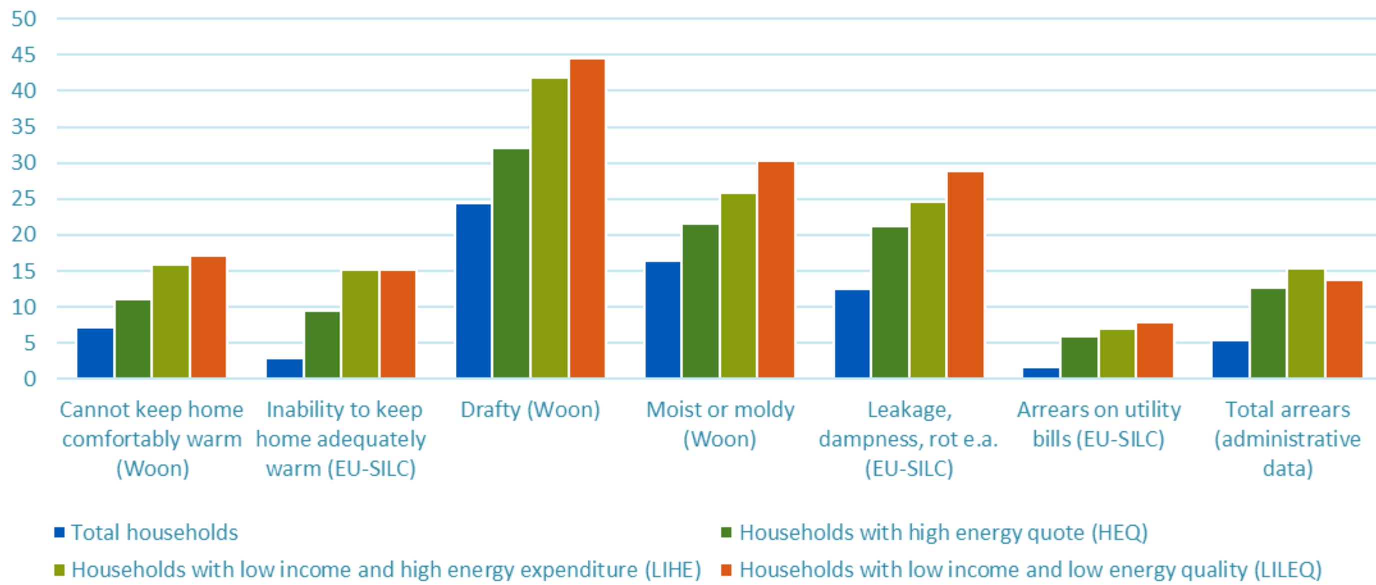

Figure 2. Households that respond positively to survey questions related to energy poverty, as a share of all households or households that are energy poor according to one of three typical indicators of energy poverty.

Estimating over- or underconsumption of energy

A final application of the model involves comparing estimated versus actual energy expenditure to infer behavioral effects. Underconsumption, where actual use falls significantly below expected levels, may indicate coping strategies in response to energy poverty (e.g., keeping the home cold). Overconsumption, by contrast, may signal inefficient behavior, poorly adjusted heating systems, or misaligned comfort preferences.

We infer over- and underconsumption by comparing actual energy expenditure with 'expected' energy consumption. We define underconsumption as using less than 75% of expected energy, and overconsumption as exceeding 125%. This method leads to 17% of both under- and overconsumption for the Netherlands in 2019. The categories of over- and underconsumption reveal behavioral adaptations, particularly among low-income groups that are more likely to under consume energy due to budget constraints, thus risking a severe form of energy poverty.

Figure 3 shows the relation between actual energy consumption of a household (x-axis) and 'expected' energy consumption at median income levels (y-axis). Expenditures are rounded to the nearest 100-fold. Each combination of rounded expenditures contains a dot which size is proportional to the number of households; the slight overlap between groups is due to rounding.

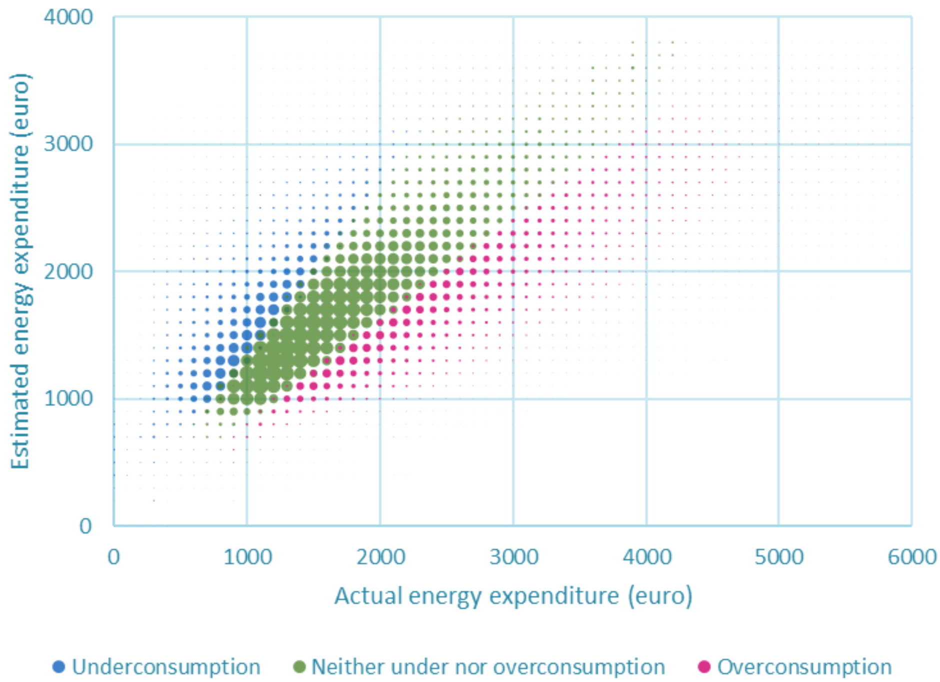

Figure 3 Relation between actual and expected energy expenditure in relation with our estimates of over- and underconsumption.

Underconsumption can have several causes, from frequently being away from home, an environmentally friendly lifestyle, freeriding on the heating efforts of neighbors in an apartment block. Yet underconsumption can also indicate energy poverty, when people turn down the thermostat to be able to pay the bills. We are especially interested in underconsumption in relation to energy poverty.

Table 8 shows the proportion of total or energy poor households with over- and underconsumption. While the total population is symmetric, with 17% for both over- and under consumption, we find that overconsumption is very common (46%) among households with a high energy quote (HEQ), still quite large (37%) for low income households with a high energy expenditure, while being smaller than average (14%) for LILEQ. For underconsumption we find that among LILEQ underconsumption is just as common as in the general population. Batenburg et al. (2024) use a similar method to derive under- and overconsumption, and find a slightly larger share of underconsumption among the energy poor (LILEQ) than among all households.

Table 8 The share of over- and underconsumption for all and energy poor households

Dutch energy poverty indicators	Underconsumption	Neither under nor overconsumption	Overconsumption
Total households for which energy poverty was measured	17	67	17
Households with high energy quote (HEQ)	1	52	46
Households with low income and high energy expenditure (LIHE)	1	62	37
Households with low income and low energy quality (LILEQ)	17	69	14
Households with low income and high energy expenditure and low energy quality (LIHELEQ)	11	62	27

4. Discussion

Our model has several weaknesses and limitations. First, the model itself is open to improvement: predictions are good but could potentially be better, and its large number of parameters make the model tedious to calculate with a risk of overfitting. Next is a more fundamental issue. Our model depends entirely on administrative data. As a result, improvements that are not formally registered—such as insulation upgrades without a new EPC or behavioral changes (e.g., reduced heating use)—are not captured. Also, changes in energy behavior are not considered. For instance, after the soaring energy prices of 2022 and 2023, many households changed their standard thermostat settings or turned off heating in bedrooms. These behavioral changes, if permanent, will result in lower inherent energy expenditure for a certain house. Yet they are not incorporated in our present model and indicator.

As a perspective, we assert that these flaws should be weighted against two large weaknesses found in alternative indicators on housing and energy. First, the standard indicator for energy efficiency (i.e. EPC labels) fails to adequately identify the homes that are associated with the highest (independent of inhabitants) energy expenditure, and thus the households most at risk (Majcen 2016, Melis et al. 2023, Few et al. 2023 Van Hove et al. 2022). Second, energy poverty indicators that are based on energy expenditure cannot make meaningful assessments of over- and underconsumption (Batenburg et al. 2024). Yet underconsumption and its related health risks are potentially a large problem associated with energy poverty. Thus, while indeed our method is flawed, we propose that it is less flawed than many of its alternatives. In addition, we propose several avenues for further research and refinement:

- Incorporation of new administrative variables: with recent expansions in administrative datasets, new variables such as the actual amount of solar electricity generated or ownership of electric vehicles, could be integrated into the model. These additions may help improve accuracy and provide further insights into household energy behavior.
- Model simplification: while the current model includes piecewise linear terms for living area to capture non-linearities, tentative tests suggest that a simpler linear specification may perform equally well when combined with other variables. Next, the interactions between variables add a lot of dimensions to the model, while the increase in predictive power (R^2) appears limited. Streamlining the model could improve computational efficiency and interpretability.
- Behavioral and dynamic data integration: although the model captures structural and socio-economic determinants of energy use, it does not account for behavioral adaptations unless these are reflected in consumption data. Integrating real-time energy use data, smart meter logs, or survey-reported comfort preferences could enhance the model's capacity to account for behavioral variation.
- Improved detection of retrofit effects: currently only administrative updates (e.g., new EPCs, solar panel or heat pump registrations) are detected as improvements in energy quality. Many insulation upgrades or heating system changes remain invisible. Linking the model to remote sensing data or retrofit subsidy records could improve the tracking of actual housing improvements.
- Cross-country applicability: while the model is tailored to Dutch data and policy frameworks, its methodology is generalizable. With appropriate adjustments for local data availability, similar models could be developed in other (EU) countries to support national energy efficiency strategies and harmonized poverty diagnostics.

5. Conclusion

This paper sets out to develop and validate a model for estimating housing energy efficiency that overcomes the limitations of conventional Energy Performance Certificates (EPCs). In doing so, we aim to contribute to accurately estimate residential energy use, better identify structurally inefficient dwellings, and more accurately target energy poverty. By combining EPC data with comprehensive administrative microdata from Statistics Netherlands (CBS), we constructed a model that yields superior predictions (in terms of explained variance, or R^2) of actual household energy consumption compared to EPC labels alone.

Our analysis demonstrates that our model captures the underlying energy performance of homes more reliably than EPCs, which are often outdated, missing, or weakly correlated with real energy use. Through the integration of structural dwelling characteristics and socio-demographic information, the model enables an estimation of energy needs which abstracts from behavioral factors. This is especially important in policy contexts, where interventions must be grounded in the technical performance of the housing stock rather than behavioral variation alone.

A key contribution of this work is the application of the model to energy poverty analysis. We offer an alternative to commonly used indictors for identifying energy poverty that are based on energy expenditure. The Low Income Low Energy Quality (LILEQ) indicator, derived from this model, has been incorporated into the national energy poverty monitoring framework established by CBS and TNO. It complements other indicators such as LIHE (Low Income High Expenditure) and HEQ (High Energy Quote), and has proven especially useful for capturing structural vulnerabilities in the housing stock. As shown in our validation using survey-based indicators (Section 4.2, Figure 1), the LEQ-based indicator (LILEQ) aligns most closely with self-reported problems related to housing quality and thermal comfort. This reinforces its value as a diagnostic tool in social and housing policy.

Another contribution lies in the detection of energy over- or underconsumption. Using model-estimated energy costs as a reference, we identify households with a structural mismatch between income and energy needs; a group that traditional expenditure-based indicators often overlook. This approach enables the detection of underconsumption, a condition where low-income households use less energy than required for adequate comfort due to financial constraints. Such households may not appear energy-poor under conventional metrics but face real material deprivation. By offering a method to isolate this condition, our model addresses an important blind spot in current monitoring practices.

In conclusion, this paper provides both a methodological contribution and a practical tool for advancing energy transition goals in the residential sector. By leveraging administrative data to estimate housing energy efficiency more accurately, we offer a foundation for more targeted, equitable, and evidence-based energy policy. As the energy transition accelerates and policy interventions become increasingly urgent, tools like the LEQ model will play a critical role in ensuring that no household is left behind.

References

Al Kez, D., Foley, A., Lowans, C., & Del Rio, D. F. (2024). Energy poverty assessment: Indicators and implications for developing and developed countries. Energy Conversion and Management, 307, 118324.

Allcott, H., & Greenstone, M. (2017). Measuring the welfare effects of residential energy efficiency programs. *NBER Working Paper No. 23386*.

Amasyali, K., & El-Gohary, N. M. (2018). A review of data-driven building energy consumption prediction studies. *Renewable and Sustainable Energy Reviews, 81*, 1192–1205.

Aydin, E., Kok, N., & Brounen, D. (2017). Energy efficiency and household behavior: The rebound effect in the residential sector. *The RAND Journal of Economics, 48*(3), 749–782.

Batenburg, A., van der Meer, L., Hopman, B., Dalla Longa, F., Geijtenbeek, L., van Middelkoop, M., Mulder, P., & Wijlhuizen, E. (2024). Energiearmoede in Nederland 2019–2023: Een overzicht van 2019 tot en met 2023 en een verdieping op onderconsumptie (TNO 2024 R10801). TNO. https://publications.tno.nl/publication/34642676/7qBbU9FJ/TNO-2024-R10801.pdf

Ben Rejeb-Mzah, I., Bedioui, M., & Larhchim, A. (2023). Machine learning to predict French residential buildings' energy performance. *Energy and Buildings, 285*, 112917.

Beijnum, B. van, van den Wijngaart, R., Luteijn, G., van der Molen, F., Geijtenbeek, L., Rovers, V., & Tigchelaar, C. (2023), Referentieverbruik warmte woningen. Achtergrondrapport, Den Haag: Planbureau voor de Leefomgeving. https://www.pbl.nl/uploads/default/downloads/pbl-2023-referentieverbruik-warmte-woningen-achtergrondrapport_5168.pdf

Civel, E., & Cruz, N. (2025). How do people read houses' energy labels? *Journal of Environmental Economics and Policy, 14(*2), 239–270.

CLO (2025). Energielabels van woningen, 2010 t/m 2024 (indicator 0556, versie 11, 18 maart 2025), www.clo.nl. Centraal Bureau voor de Statistiek (CBS), Den Haag; PBL Planbureau voor de Leefomgeving, Den Haag; RIVM Rijksinstituut voor Volksgezondheid en Milieu, Bilthoven; en Wageningen University and Research, Wageningen.

Cui, X., Lee, M., Koo, C., & Hong, T. (2024). Energy consumption prediction and household feature analysis for different residential building types using machine learning and SHAP. *Energy and Buildings, 309,* 113997.

Department for Business & Energy and Industrial Strategy. Fuel poverty methodology handbook: Low Income Low Energy Efficiency (LILEE). London, United Kingdom; 2022. <https://www.gov.uk/government/publications/fuel-poverty-statistics-methodology-handbook>.

European Commission, 2025, "Directive (EU) 2024/1275 of the European Parliament and of the Council of 24 April 2024 on the energy performance of buildings (recast) (Text with EEA relevance)". European Data Portalaccessdate=26 May 2025. 24 April 2024.

Few, J., Manouseli, D., McKenna, E., Pullinger, M., Zapata-Webborn, E., & Oreszczyn, T. (2023). The over-prediction of energy use by EPCs in Great Britain. *Energy and Buildings, 288*, 113024.

Fowlie, M., Greenstone, M., & Wolfram, C. (2018). Do energy efficiency investments deliver? *The Quarterly Journal of Economics, 133(*3), 1597–1644.

Harputlugil, T., & de Wilde, P. (2021). The interaction between humans and buildings for energy efficiency. *Energy Research & Social Science, 71*, 101828.

Ishwaran, H. (2007). Variable importance in binary regression trees and forests.

Karpinska, L., & Smiech, S. (2023). Multiple faces of poverty: Exploring housing-costs-induced energy poverty in Central and Eastern Europe. *Energy Research & Social Science, 105*, 103273.

Kattenberg, L., Eichholtz, P., & Kok, N. (2022). The efficacy of energy efficiency: Measuring the returns to home insulation. *Maastricht University Research Paper*.

Majcen, D. (2016). Predicting energy consumption and savings in the housing stock: A performance gap analysis in the Netherlands. *A+BE | Architecture and the Built Environment, 6(*4), 1–224.

Majcen, D. Itard, L., & Visscher, H. (2013). Theoretical vs. actual energy consumption of labelled dwellings in the Netherlands: discrepancies and policy implications, *Energy Policy*, 54 (2013), pp. 125-136.

Majcen, D. Itard, L., & Visscher, H. (2016). Actual heating energy savings in thermally renovated Dutch dwellings, *Energy Policy*, 98 (2016), pp. 82-92.

Melis, T., Farrell, N., & Curtis, J. (2023). How well do building energy performance certificates predict heat loss? *Energy Efficiency, 16,* 74.

Meyer, S., Laurence, H., & Bartiaux, F. (2018). The 'sticky' effect of energy poverty: Biographical experiences of energy poor households in Belgium. Energy Research & Social Science, 40, 119–129. https://doi.org/10.1016/j.erss.2018.01.015

Mulder, P., Dalla Longa, F., & Straver, K. (2023). Energy poverty in the Netherlands at the national and local level: A multi-dimensional spatial analysis. Energy Research & Social Science, 96, 102892. https://doi.org/10.1016/j.erss.2022.102892

Myers, E., Puller, S. L., & West, J. (2022). Mandatory energy efficiency disclosure in housing markets. *American Economic Journal: Economic Policy, 14*(4), 453–487.

Papineau, M., Rivers, N., & Yassin, K. (2025). Household benefits from energy efficiency retrofits: Implications for net zero housing policy. *Energy Economics, 143,* 108245.

Peñasco, C., & Díaz Anadón, L. (2023). Assessing the effectiveness of energy efficiency measures in the residential sector. *Energy Economics, 117*, 106435.

Rademaekers, K., Yearwood, J., Ferreira, A., Pye, S. T., Hamilton, I., Agnolucci, P., ... & Anisimova, N. (2016). Selecting indicators to measure energy poverty.

Rezvy, S., Zebin, T., & Mohammed, R. (2023). Building energy usage predictions using machine learning and explainable AI methods. *Energy and AI, 10*, 100183.

Semple, T., Rodrigues, L., Harvey, J., et al. (2024). An empirical critique of the low income low energy efficiency approach. *Energy Policy, 186*, 114014.

Statistics Netherlands (2024). Methodedocument monitor energiearmoede 2019-2022. https://www.cbs.nl/nl-nl/longread/aanvullende-statistische-diensten/2024/methodedocument-monitor-energiearmoede-2019-2022

Thomson, H., Snell, C., & Bouzarovski, S. (2017). Health, well-being and energy poverty in Europe: A comparative study of 32 European countries. International Journal of Environmental Research and Public Health, 14(6), 584. https://doi.org/10.3390/ijerph14060584

Van den Wijngaart, R., & Van Polen, S. (2020). Bepaling energiebesparing door isolatie van woningen in de Startanalyse 2020, PBL-publicatienummer 4284.

Van Hove, M. Y. C., et al. (2022). Large-scale statistical analysis and modelling of real and regulatory energy use. *Building Research & Information, 51*(2), 203–222.