

Article

Model Validity and Transferability Informing Behavioral Energy Policies

Areti Kontogianni ^{1,*}, Dimitris Damigos ², Michail Skourtos ³, Christos Tourkolias ⁴, Eleanor Denny ⁵, Ibon Galarraga ⁶, Steffen Kallbekken ⁷ and Edin Lakić ⁸

¹ Department of Mechanical Engineering, University of Western Macedonia, 50100 Kozani, Greece

² School of Mining and Metallurgical Engineering, National Technical University of Athens, 15780 Zografou Campus, Greece; damigos@metal.ntua.gr

³ Department of Economics, College of Business Administration, University of Kuwait, Kuwait 13060, Kuwait; michail.skourtos@ku.edu.kw

⁴ Center for Renewable Energy Sources, 19009 Pikermi, Greece; ctourkolias@teemail.gr

⁵ Department of Economics, Trinity College, D02 PN40 Dublin, Ireland; dennye@tcd.ie

⁶ BC3 Basque Centre for Climate Change, 48940 Leioa, Spain; ibon.galarraga@bc3research.org

⁷ CICERO Center for International Climate Research, P.O. Box 1129 Blindern, 0349 Oslo, Norway; steffen.kallbekken@cicero.oslo.no

⁸ Faculty of Electrical Engineering, University of Ljubljana, 1000 Ljubljana, Slovenia; edin.lakic@fe.uni-lj.si

* Correspondence: akontogianni@uowm.gr



Citation: Kontogianni, A.; Damigos, D.; Skourtos, M.; Tourkolias, C.; Denny, E.; Galarraga, I.; Kallbekken, S.; Lakić, E. Model Validity and Transferability Informing Behavioral Energy Policies. *Energies* **2021**, *14*, 3122. <https://doi.org/10.3390/en14113122>

Academic Editors: Carlos Henggeler Antunes and Christos Vlachokostas

Received: 24 April 2021

Accepted: 25 May 2021

Published: 27 May 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: A number of microeconomic choice models are currently applied to demonstrate systematic biases in energy consumer behavior. The models highlight the hidden potential of energy savings from policies that target the so-called behavioral anomalies. Nevertheless, whether these patterns are repeatable or not is not clear, because the efforts to determine the transferability or generalizability of these models are practically nonexistent. This paper uses a unique collection of empirical data from five EU countries collected within the CONSEED project to refine and develop further the standard consumer decision model, validate it for policy purposes, and elaborate on its transferability between countries. The pooled samples allow for a more reliable investigation of the relative importance of the factors influencing consumers' attitudes and beliefs towards energy investment decisions. Based on the statistical tests conducted to evaluate the "transferability" of the pooled models (i.e., the possibility of creating a "universal" model of EE from the pooled model), it can be argued that the models are transferable in specific cases since attitudinal factors and demographic characteristics play a significant role. Although the pooled models are validated, any extrapolation of the above-mentioned findings to specific populations in terms of "space" (i.e., country) and "target" (e.g., sectors and technologies) should be approached with caution from a policy perspective.

Keywords: model validation; behavioral models; energy efficiency

1. Introduction

Current strategies to address climate change share a fundamental assumption: Containment of greenhouse gases hinges essentially on the success of behavioral changes in energy use and investment decisions [1–3]. The low uptake of energy-efficient appliances can be, at least partially, explained by informational failures (such as imperfect and asymmetric information and myopia) and behavioral biases related, for instance, to social norms, decision-making heuristics, and inattention, e.g., [4–9]. Thus, it is expected that more informed, forward-looking electricity consumers would contribute to energy reduction worldwide, e.g., [10–14]. The confinement measures born out of the recent pandemic have eased energy use [15] and brought the need for restructuring the energy sector to the fore. At the same time, the measures vividly showed what can be achieved when individuals are persuaded to follow new behavioral norms [16].

Several studies have used multiple regression models to identify patterns between market and non-market barriers and energy-efficiency investment decisions, e.g., [4,5,8,9,12,17–20]. Nevertheless, whether these patterns are repeatable or not is not clear, because the efforts to determine the transferability or generalizability of these models are practically nonexistent. The issue of transferability is of great interest in ecological models (e.g., [21–23]), models of air pollutants (e.g., [24–26]), travel forecasting models (e.g., [27–29]), and especially in non-market valuation (an approach known as “benefit function transfer,” e.g., [30–34]). Concerning energy studies, the aspect of transferability has been tested in predictive models of energy customer data analytics [35], building energy consumption data models [36], and residential energy demand and city-scale electricity usage models [37,38]. In addition, Warren [39] developed a multicriteria decision-making framework to identify whether demand-side management (DSM) policies are transferable and Bößner et al. [40] investigated the transfer of renewable energy support policies between countries.

Bearing in mind the above-mentioned remarks, the objective of the present paper is to contribute to a behaviorally informed energy-saving strategy by filling the existing gap in the literature regarding the transferability potential of energy consumer decision-making models. As a first step, datasets were collected within the CONSEED project (<https://www.conseedproject.eu>, accessed on 30 March 2021) in a number of discrete choice experiments, along with surveys and field trials. The target was different consumer groups, spatial locations, and product categories, e.g., [12,20,41–43]. The data were examined and cleaned to be made suitable to pool for econometric analysis. This resulted in two homogenized datasets, as detailed in Section 3. The two datasets were used to validate the theoretical and econometric models developed prior to the survey (for details see [44]) Specifically, the pooled models aim to determine the relative importance of (a) market and non-market biases (e.g., mistrust of energy labels, “warm glow” motives associated with the perceived negative environmental and societal consequences of energy consumption, etc.) and (b) socio-demographic characteristics that account for agent heterogeneity in energy investment decisions. These two datasets are used to evaluate the “transferability” of the pooled models between countries and technologies, following the validation methodology discussed in Section 2.2, an underappreciated aspect of statistical validation [22]. Hence, the findings presented here are related to the following two hypotheses:

Proposition 1. *The pooled model is valid and, therefore, provides a reliable view of the market and non-market barriers affecting consumer energy efficiency decisions.*

Proposition 2. *The pooled model is “transferable” between countries and technologies; that is, the identified relationships between consumer energy-investment decisions, socioeconomic factors, and market and non-market barriers are general and not idiosyncratic to a limited set of conditions.*

The first proposition examines the validity of the model in terms of goodness of fit and discrimination using appropriate criteria (i.e., Pearson’s χ^2 goodness-of-fit statistic and C-statistic). The null hypothesis is that the pooled models are valid. The second proposition tests the transferability potential of the validated models between different countries, technologies, sectors, etc. The evaluation is performed using Pearson’s χ^2 goodness-of-fit and C-statistic indicators, as well as the Wald test regarding the equality of the coefficients across the models. Previous research has shown that energy efficiency-related attitudes and behavior are affected by sociodemographic characteristics [5,45], attitudinal characteristics [45], and market and non-market failures (e.g., asymmetric/imperfect information, misplaced motivations, biased perceptions on the product energy consumption, etc.) [13,45–51]. The above-mentioned factors differ across countries, consumer goods, etc., and thus, in this case, the null hypothesis is that the models are not “transferable” in principle.

The paper is structured as follows: Section 2 provides the methodological overview. Section 3 presents and discusses the main results drawn from the analysis. We conclude in Section 4 with thoughts on the potential of our model(s) to feed into cross-national European energy policy design.

2. Materials and Methods

2.1. Sample Pooling

The CONSEED partners initially set out to design surveys, discrete choice experiments, and field trials in a harmonized way to facilitate subsequent data pooling. However, the considerable heterogeneity in field trials and discrete choice experiments led to inconsistency in variables and contexts that prevented the pooling of these datasets. For example, although the results of the property and appliance field trials are comparable in terms of their treatment effects, it is not feasible or desirable to combine these two datasets consistently for ex-post analysis and verification purposes due to the very different characteristics of each dataset. Furthermore, although each discrete choice experiment included EE as an attribute, all remaining attributes differed by the technology explored, as did the experimental designs and treatment methods. Given the heterogeneity of the field trials and discrete choice experiments, pooling would not produce a workable dataset.

Of the available data types, only the consumer surveys addressed to households and firms (including farmers) are suitable for pooling and analysis. Despite considerable differences in locations (Greece, Ireland, Spain, Norway, and Slovenia), technologies (property, appliances, transport, and machinery), and sectors (households, services, agriculture, and industry), CONSEED designed a “core” set of survey questions that were applied in all surveys (see, for example, Table S1: Common survey questions). Often, this required minor adjustments depending on the sample and context. For example, when asked about the importance of technology attributes, the set of six attributes always included energy/fuel efficiency/consumption and price, but partners had discretion in choosing four additional sample-specific attributes. In addition to this, partners had discretion in asking supplementary non-core questions in their surveys. The resulting “pooled” dataset incorporates all core survey questions only, covering households and firms.

The “households” pooled sample included 3016 observations from Norway (1093 or 36.2%), Ireland (501 or 16.6%), Spain (500 or 16.6%), Greece (496 or 16.4%), and Slovenia (426 or 14.1%). Further, 1093 observations (36.2%) concerned transportation goods, i.e., cars; 996 observations (33%) concerned household appliances, i.e., refrigerators (496 observations) and washing machines (500 observations); and 927 (30.7%) observations concerned properties. As far as the “firms” pooled sample is concerned, it included 794 observations in total. More specifically, 492 observations (62%) derived from Ireland, 200 observations (25.2%) from Spain, and 102 observations (12.8%) from Greece. As regards the technological subsamples, 302 observations (38%) involved heating and cooling appliances in the services sector, 316 observations (39.8%) concerned machinery in the agricultural sector, and 176 observations (22.2%) concerned heating controls in the properties sector.

2.2. Model Validation

Model validation is a statistical concept that hints at the model’s degree of generalizability and implies that the model chosen for measuring the theoretical concepts and relationships accurately represents the real-world phenomena [52]. Specifically, the focus of this paper is on assessing the performance of the model on predicting the probability of the positive event for out-of-the-sample subjects [53–55].

There are two main statistical approaches to model validation, namely, external validation (the model is tested on a new sample) and internal validation (the model is tested on a subsample of the original sample). Although external validation appears to be a rigorous validation approach, it has certain disadvantages. For instance, the collection of a new sample that will only be used in the validation of the model can be rather costly. Further, even if monetary resources are available, the new sample must be sufficiently representative of the original sample, otherwise the validation process can provide misleading results, either pessimistic or optimistic. For these reasons and due to budget constraints, the internal validation approach was adopted in this case.

The most accredited internal validation methods are data-splitting, repeated data-splitting, jackknife, and bootstrapping [55]. In all these methods, a subsample of observa-

tions is excluded from the analysis, a model based on the remaining subjects is developed, and then it is tested on the originally excluded subjects. In this study, the repeated data splitting approach was chosen, for reasons detailed in [55], and hence, the available data were split into two portions. The first portion was used to fit the model (i.e., fitting sample) and the second was used to assess its performance (i.e., validation sample). Usually, the portion of observations reserved for validation is always less than $1/2$ and in the range of $1/4$ to $1/3$ [56], and a priori the performance of the validation model is inferior to that of the fitting sample [55]. Fitting and validation samples in our cases were defined from original and pooled datasets.

The model validation process of the pooled samples involves the following steps described in [55], after modifications:

Step 1: Data splitting. The pooled sample is randomly split into the fitting (70%) and validation samples (30%, to have an adequate number of observations).

Step 2: Model fitting. The model's coefficients, its overall significance, and the partial significance of each variable are estimated on the fitting sample.

Step 3: Event probability estimation. The coefficients from Step 2 are used to estimate the probability of a positive outcome for each of the subjects in the fitting and the validation samples.

Step 4: Computation of performance measures. The fitting and validation samples are tested against the following statistics:

- C-statistic (a measure of predictive power);
- Pearson's χ^2 goodness-of-fit statistic.

The C-statistic is a measure of goodness of fit for binary outcome regression models and it is equal to the area under the receiver operating characteristic (ROC) curve and ranges from 0.5 to 1. As a general rule [53],

if $C \geq 0.9$, the model is considered to have outstanding discrimination;

if $0.8 \leq C < 0.9$, the model is considered to have excellent discrimination;

if $0.7 \leq C < 0.8$, the model is considered to have acceptable discrimination;

if $C = 0.5$, the model has no discrimination;

if $C < 0.5$, the model has negative discrimination, i.e., it is worse than random.

Pearson's χ^2 goodness-of-fit statistic, as a measure of calibration, examines the null hypothesis that the logistic regression model used is the correct model [57].

Step 5: Iterations. The above-described procedure is repeated 100 times. Each iteration is based on a different split of the original data and results in different model coefficients, significance levels, and performance values.

Step 6: Results. After the 100 repetitions, the distribution of each of the performance measures is provided for the fitting and validation samples. As mentioned, in general, a reduction in the magnitude of discrimination and calibration performance is expected, but if the drop is too large, the model does not validate outside the fitting sample [55]. To assess model validity, the equality of the two ROC areas obtained from fitting and validation samples is tested using an algorithm suggested by [58] and the critical values for the Pearson's χ^2 test are examined for the distributions of the 100 repetitions. The distributions of the goodness-of-fit estimates of fitting and the validation samples should be averaged around the same values. If this does not happen, the model cannot be validated because of its internal instability [55].

Step 7: Evaluation of "transferability". After having validated the model for the pooled sample, the same model is fitted to different subsamples constructed by country and technology to evaluate the model "transferability" between countries and technologies using the same performance indicators. In this case, however, the Wald test is also used [59] to test the equality of the coefficients across the models. Given that the pooled datasets come from surveys of households and firms, the above-mentioned process is implemented separately for the two groups of interest.

3. Results and Discussion

3.1. Pooled Model Validation—Households

The pooled sample included observations from Greece (appliances), Ireland (properties), Norway (transport), Slovenia (properties), and Spain (appliances). To validate the pooled sample model, the missing observations were dropped to correctly compare the curves based on the non-missing data. The final dataset contained 1896 observations out of 3016 (62.8%). Several models were run using diverse sets of predictors, which were initially selected by considering the estimated binary response models. The dependent variable in all models was “energy efficiency is a very important attribute in the purchasing decision” (coded “1” for “Yes” and “0” otherwise). Table 1 presents the marginal effects (at the mean) of the explanatory variables selected on the probability that respondents included in the pooled sample valued energy efficiency as a “very important” attribute.

Table 1. Variables affecting households’ perceptions of energy efficiency—probit marginal effects (pooled sample).

Variables	Pooled Sample
	Margins
Attitudes Towards EE	
Environmental impact reduction (=1 if strongly disagree)	0.043 ** (0.020)
EE appliances/cars less reliable (=1 if strongly disagree)	−0.056 *** (0.015)
Chance on new technologies (=1 if strongly disagree)	0.081 *** (0.019)
Lack access to loans, etc. (=1 if strongly disagree)	0.030 *** (0.011)
Aware of energy prices (=1 if strongly disagree)	0.053 *** (0.015)
Understand EE cost reduction (=1 if strongly disagree)	0.055 *** (0.016)
Perception of Existing Label	
Energy labels affect my choices (=1 if strongly disagree)	0.111 *** (0.017)
Environmental Awareness	
Concerned about the environment (=1 if not concerned)	0.041 *** (0.015)
Sociodemographics	
Gender (1: male; 2: female)	0.106 *** (0.024)
Pseudo-R ²	0.109
Pearson’s χ^2	0.106
Observations	1896

Note: Standard errors in parentheses *** $p < 0.01$, ** $p < 0.05$.

According to the model, those who strongly agreed that purchasing a more energy-efficient appliance/car/property would reduce the impact of their household compared to those who slightly agreed were 4.3% more likely to consider energy efficiency “very important,” keeping everything else constant. Further, the probability of considering energy efficiency “very important” was 5.6% lower for those believing that energy-efficient appliances/cars/properties are less reliable than those who did not share the same beliefs. Similarly, those who believed that new technologies would help to reduce energy consumption, were aware of energy prices, were “energy literate” (i.e., they understood energy-efficiency savings), and were concerned about the environment had a higher probability of considering energy efficiency a very important attribute than those who believed the opposite (by 8.1%, 5.3%, 5.5%, and 4.1%, respectively). Credit constraint respondents also believed that energy efficiency is a very important attribute (the probability was 3%

higher). Further, women were on average 10.6% more likely to value the energy-efficiency attribute as very important, compared to men. Finally, and most importantly, the respondents who claimed that energy labels would affect their choice were 11% more likely to value the energy efficiency as “very important”.

The pooled model was tested for validity using the process described in Section 2.2. The results for the distribution of each of the performance measures for the fitting and validation samples are summarized in Table 2. According to the ROC comparisons column, the chi-squared test yielded an average significance probability of 0.502, suggesting that there was no significant difference between the two areas of the fitting and validation models. The null hypothesis, i.e., the areas under the ROC curves are equal, was rejected in 9% of the iterations. Thus, it can be argued that the pooled model discriminates very well. The discriminative ability of the model implies that a randomly selected respondent who believes that energy efficiency is very important will have a higher predicted probability of having this outcome occur, compared to a randomly selected respondent who does not believe the same. In other words, it denotes that the model allows us to discriminate between low and high energy-efficiency importance observations.

Table 2. Distribution of performance measures—household pooled model validation.

Parameter	Values		
	ROC Comparisons Prob > χ^2	Pearson’s χ^2 —Fitting	Pearson’s χ^2 —Validation
Mean	0.502	0.169	0.264
Std. Dev.	0.316	0.08	0.141
Percentiles			
1%	0.006	0.043	0.007
5%	0.038	0.062	0.049
10%	0.053	0.077	0.085
25%	0.195	0.106	0.159
50%	0.528	0.160	0.260
75%	0.783	0.214	0.352
90%	0.945	0.26	0.452
95%	0.97	0.343	0.51
99%	0.993	0.402	0.63

Further, as regards the goodness of fit of the two models, according to Pearson’s χ^2 results, the fitting model was rejected in only 2% of the iterations and the validation model in approximately 5% of the iterations. This is also evident in Figure 1, according to which the models could not be rejected in almost any iteration.

3.2. Transferability Evaluation—Households Models

To explore the “transferability” of the pooled model to different countries and technologies (i.e., to examine whether it is possible to create a “universal” model of energy efficiency), eight different models were run for the five countries and three technologies and the results were compared based on three criteria: the estimated ROC curve areas, the goodness of fit of the models according to the Pearson’s χ^2 test, and the Wald test regarding the equality of the coefficients. The dependent variable in all the models was “energy efficiency is a very important attribute in the purchasing decision” (coded “1” for “Yes” and “0” otherwise).

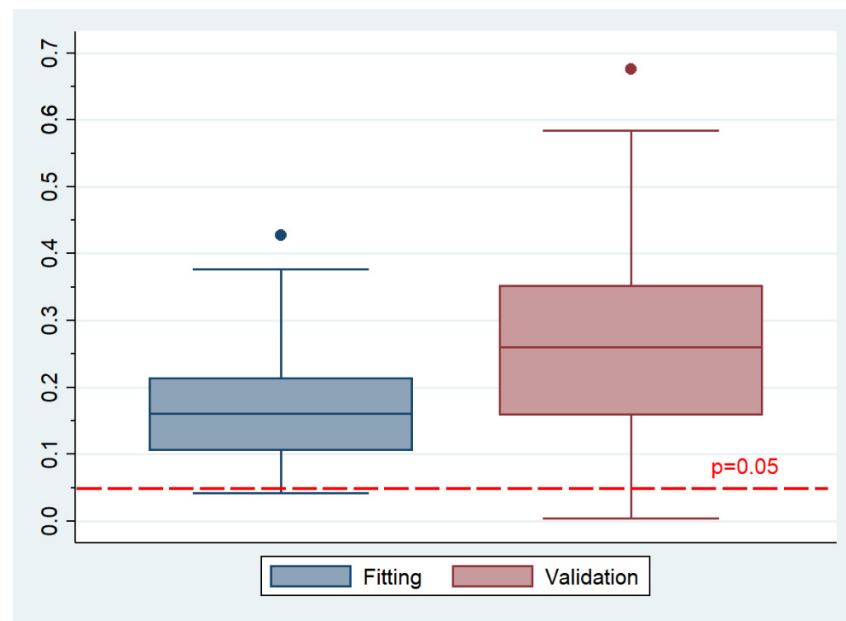


Figure 1. Pearson's χ^2 —fitting vs. validation model distributions (households).

3.2.1. Country Transferability

The results for the five national subsamples are given in Tables 3 and 4. In general, the calibration criterion was satisfied for all countries (yet only marginally for the Irish subsample). According to ROC comparisons, the chi-squared test yielded an average significance probability of 0.062, suggesting that there was no significant difference between the ROC areas of the five national models.

Nevertheless, the C-statistic criterion was fulfilled for three out of the five countries (Table 4 and Figure 2), namely, Greece, Ireland, and Slovenia. The models for Norway and Spain, although marginally, did not present acceptable discrimination.

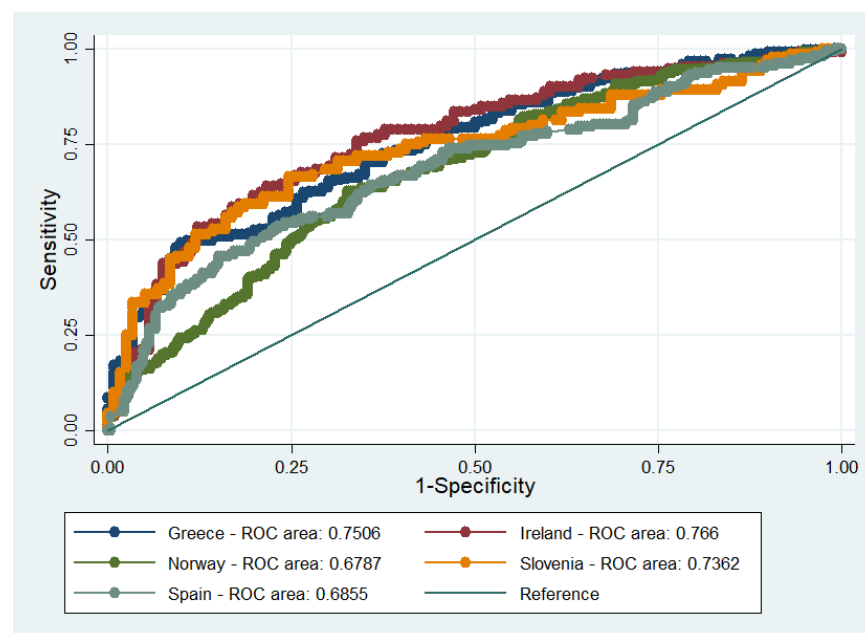


Figure 2. ROC curves and C-statistic estimates for the national models (households).

Table 3. Variables affecting households' perceptions of energy efficiency—probit marginal effects (country subsamples).

Variables	Greece	Ireland	Norway	Slovenia	Spain
	Margins	Margins	Margins	Margins	Margins
Attitudes Towards EE					
Environmental impact reduction (=1 if strongly disagree)	0.031 (0.041)	0.104 * (0.056)	0.043 (0.032)	0.034 (0.047)	0.072 (0.053)
EE appliances/cars less reliable (=1 if strongly disagree)	−0.082 *** (0.027)	−0.047 (0.036)	−0.014 (0.025)	−0.058 (0.043)	−0.031 (0.040)
Chance on new technologies (=1 if strongly disagree)	0.070 (0.044)	0.084 (0.052)	0.058 * (0.032)	0.081 * (0.049)	0.092 ** (0.046)
Lack access to loans, etc. (=1 if strongly disagree)	0.024 (0.026)	0.032 (0.035)	0.007 (0.021)	−0.047 (0.032)	−0.027 (0.030)
Aware of energy prices (=1 if strongly disagree)	0.089 *** (0.033)	0.047 (0.054)	0.087 ** (0.035)	0.034 (0.049)	0.022 (0.032)
Understand EE cost reduction (=1 if strongly disagree)	0.005 (0.032)	0.033 (0.048)	0.071 ** (0.030)	0.110 * (0.059)	−0.038 (0.038)
Perception of Existing Label					
Energy labels affect my choices (=1 if strongly disagree)	0.099 *** (0.031)	0.158 *** (0.047)	0.055 * (0.031)	0.121 ** (0.047)	0.152 *** (0.041)
Environmental Awareness					
Concerned about the environment (=1 if not concerned)	0.055 ** (0.027)	0.097 ** (0.048)	0.014 (0.026)	0.072 * (0.039)	0.023 (0.038)
Sociodemographics					
Gender (1: male; 2: female)	0.073 (0.046)	0.095 (0.069)	0.141 *** (0.044)	0.134 * (0.071)	0.061 (0.052)
Pseudo-R ²	0.143	0.145	0.074	0.124	0.078
Pearson's χ^2 p-values	0.160	0.052	0.414	0.255	0.229
Observations	423	239	582	258	394

Note: Standard errors in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 4. Comparison of the ROC areas—country subsample models (households).

Country	Obs.	ROC Area	Std. Err.	Asymptotic Normal [95% Conf. Interv]	
Greece	423	0.751	0.025	0.701	0.800
Ireland	239	0.766	0.031	0.705	0.827
Norway	582	0.679	0.022	0.635	0.722
Slovenia	258	0.736	0.031	0.675	0.797
Spain	394	0.686	0.028	0.631	0.740

Since the value and the statistical significance of the coefficients across the pooled and the national sub-models differed, the Wald test was implemented to test whether this difference was statistically significant. The results are given in Table 5. The null hypothesis that all the coefficients in the examined models are equal was rejected between the pooled model and the models for Spain and Norway, as well as between the models for Spain and Norway. These results are consistent with the remarks made about the discrimination of these two models. All in all, it could be argued that the model is not transferable across all countries mainly because the importance of energy efficiency varied between the samples.

3.2.2. Technology Transferability

The results of the technological subsamples are presented in Tables 6 and 7. The null hypothesis that the regression model used is the correct model could not be rejected at a 5% level (it was rejected at a 10% level for the properties model). Based on Pearson's χ^2 test, it was suggested that there is a significant difference between the ROC areas of the

three models ($p = 0.034$). The test was rejected due to the transport model, which also had unacceptable discrimination given that the C-statistic was below 0.7 (Table 7 and Figure 3).

Table 5. Wald test results for the pooled and national subsample models (households).

Country Difference	Prob > χ^2	Country Difference	Prob > χ^2
Pooled–Greece	0.6309	Greece–Spain	0.3868
Pooled–Ireland	0.8729	Ireland–Norway	0.2851
Pooled–Norway	0.0008	Ireland–Slovenia	0.8248
Pooled–Slovenia	0.4483	Ireland–Spain	0.7184
Pooled–Spain	0.0137	Norway–Slovenia	0.6733
Greece–Ireland	0.9371	Norway–Spain	0.0298
Greece–Norway	0.1386	Slovenia–Spain	0.4350
Greece–Slovenia	0.7912		

Table 6. Variables affecting households' perceptions of energy efficiency—probit marginal effects (technology subsamples).

Variables	Appliances	Properties	Transport
	Margins	Margins	Margins
Environmental impact reduction (=1 if strongly disagree)	0.040 (0.036)	0.059 * (0.035)	0.043 (0.032)
EE appliances/cars less reliable (=1 if strongly disagree)	−0.079 *** (0.025)	−0.059 ** (0.027)	−0.014 (0.025)
Chance on new technologies (=1 if strongly disagree)	0.117 *** (0.034)	0.086 ** (0.035)	0.058 * (0.032)
Lack access to loans, etc. (=1 if strongly disagree)	0.050 *** (0.019)	−0.012 (0.023)	0.007 (0.021)
Aware of energy prices (=1 if strongly disagree)	0.081 *** (0.024)	0.026 (0.035)	0.087 ** (0.035)
Understand EE cost reduction (=1 if strongly disagree)	0.020 (0.026)	0.072 ** (0.035)	0.071 ** (0.030)
Energy labels affect my choices (=1 if strongly disagree)	0.142 *** (0.027)	0.130 *** (0.032)	0.055 * (0.031)
Concerned about the environment (=1 if not concerned)	0.044 * (0.024)	0.076 *** (0.029)	0.014 (0.026)
Gender (1: male; 2: female)	0.065 * (0.037)	0.121 ** (0.049)	0.141 *** (0.044)
Pseudo-R ²	0.143	0.124	0.074
Pearson's χ^2	0.179	0.093	0.414
Observations	817	497	582

Note: Standard errors in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 7. Comparison of the ROC areas—technology subsample models (households).

Technology	Obs	ROC Area	Std. Err.	Asymptotic Normal (95% Conf. Interval)	
Appliances	817	0.747	0.017	0.714	0.78
Properties	497	0.745	0.022	0.701	0.788
Transport	582	0.679	0.022	0.635	0.722

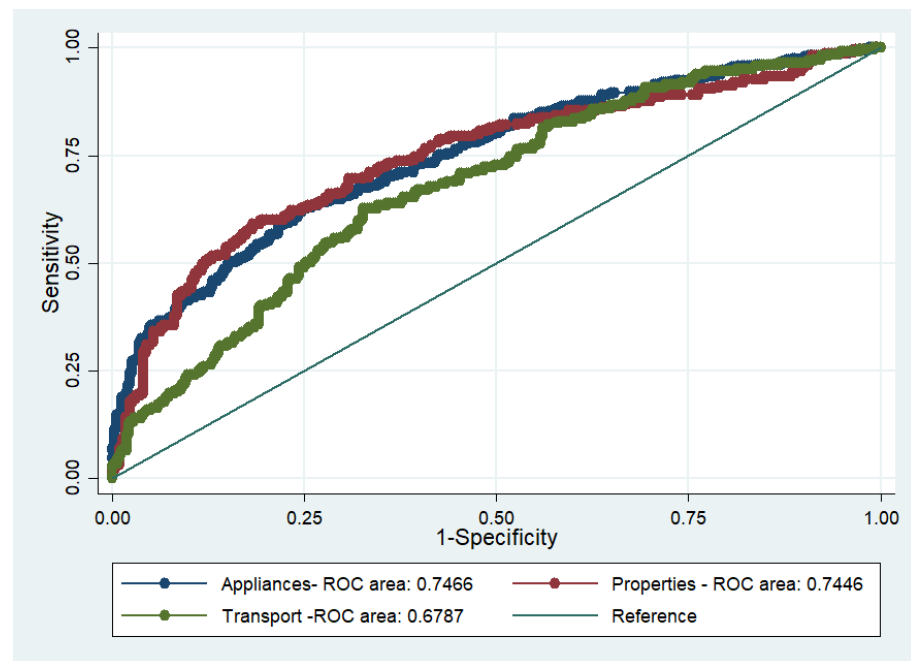


Figure 3. ROC curves and C-statistic estimates for the technological models (households).

Finally, the “transferability” of the model was examined using the Wald test, the results of which are summarized in Table 8. The null hypothesis that all the coefficients in the examined models are equal was rejected between the pooled model and the transport model, and between the transport model and the appliances model. Again, the Wald test results coincided with the results of discrimination for the transport model. Overall, it could be argued that the model is transferable as regards the appliances and properties models, but not the transport model.

Table 8. Wald test results for the pooled and technological subsample models (households).

Country Difference	Prob > χ^2
Pooled–Appliances	0.1284
Pooled–Properties	0.4709
Pooled–Transport	0.0008
Appliances–Properties	0.4161
Appliances–Transport	0.0171
Properties–Transport	0.1954

3.3. Pooled Model Validation—FIRMS

The pooled sample for the firms included observations from Greece (appliances in services sector), Spain (appliances in services sector), and Ireland (properties and services sector and machinery in agricultural sector). The missing observations were dropped from the dataset prior to validating the pooled sample model to correctly compare the ROC curves. The final dataset contained 555 observations out of 794 (70%). Several models were run using diverse sets of predictors, which were initially selected by considering the estimated binary response models. Similarly, the dependent variable in all the models was “energy efficiency is a very important attribute in the purchasing decision” (coded “1” for “Yes” and “0” otherwise). Table 9 presents the marginal effect of the factors affecting the probability of considering energy efficiency a “very important” attribute in the pooled sample.

Table 9. Variables affecting firms' perceptions of energy efficiency—probit marginal effects (pooled sample).

Variables	Pooled Sample
	Margins
Attitudes Towards EE	
Environmental impact reduction (=1 if strongly disagree)	0.067 ** (0.028)
Chance on new technologies (=1 if strongly disagree)	0.078 *** (0.030)
Lack access to loans, etc. (=1 if strongly disagree)	0.046 ** (0.019)
Understand EE cost reduction (=1 if strongly disagree)	0.100 *** (0.025)
Perception of Existing Label	
Energy labels affect my choices (=1 if strongly disagree)	0.047 * (0.026)
Pseudo-R ²	0.085
Pearson's χ^2	0.536
Observations	555

Note: Standard errors in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

According to the model, the marginal effect at the means (MEM) for those who strongly agreed that acquiring more energy-efficient equipment would reduce their firm's environmental impact was 6.7% (i.e., on average they were about 7% more likely to value energy efficiency as "very important"). The MEMs for those who were willing to implement new energy-efficient technologies and had a good understanding of energy savings were 7.8% and 10.0% more likely to rate energy efficiency as "very important," respectively. Further, respondents who claimed that the lack of access to loans acts as a barrier to energy investments also believed that energy efficiency is a very important attribute (MEM: 4.6%). Finally, interviewees who claimed that energy labels would affect their choice were 4.7% more likely to value the energy efficiency attribute as "very Important".

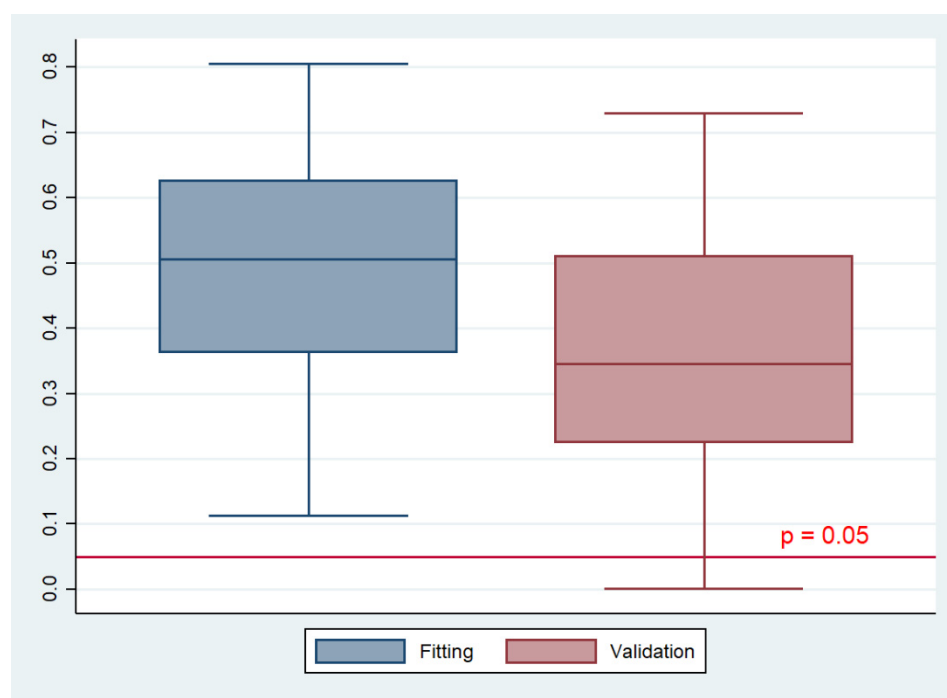
The pooled model was tested for validity using the process described in Section 2.2. The results for the distribution of each of the performance measures for the fitting and validation samples are summarized in Table 10. Based on the ROC comparisons results, the chi-square test yielded an average significance probability of 0.565, which suggests that there was no statistically significant difference between the fitting and validation model ROC areas. The null hypothesis, i.e., the areas under the ROC curves are equal, was rejected in 3% of the iterations. Thus, it can be argued that the model discriminates very well. With regards to the goodness of fit of the two models, according to the Pearson's χ^2 results the fitting model was not rejected in any of the iterations and the validation model was rejected in 2% of the iterations. Furthermore, as illustrated in Figure 4, the models could not be rejected practically in any of the iterations.

3.4. Transferability Evaluation—Firm Models

The "transferability" of the firms' pooled model was also tested for different countries, technologies, and sectors. In total, eight different models were run for the three countries, three technologies, and two sectors, which were compared using the three criteria, i.e., the estimated area of ROC curves, the goodness-of-fit of the models according to the Pearson's χ^2 test, and the Wald test regarding the equality of the coefficients. The dependent variable in all the models was "energy efficiency is a very important attribute in the purchasing decision" (coded "1" for "Yes" and "0" otherwise).

Table 10. Distribution of performance measures—firm pooled model validation.

Parameter	Values		
	ROC Comparisons Prob > χ^2	Pearson's χ^2 —Fitting	Pearson's χ^2 —Validation
Mean	0.565	0.491	0.364
Std. Dev.	0.264	0.166	0.188
Percentiles			
1%	0.002	0.123	0.015
5%	0.093	0.222	0.078
10%	0.151	0.266	0.115
25%	0.367	0.363	0.225
50%	0.612	0.506	0.345
75%	0.782	0.626	0.510
90%	0.913	0.683	0.639
95%	0.959	0.763	0.685
99%	0.976	0.804	0.724

**Figure 4.** Pearson's χ^2 —fitting vs. validation model distributions.

3.4.1. Country Transferability

The results for the three national subsamples are given in Tables 11 and 12. The calibration criterion was generally satisfied for all countries. As far as the C-statistic criterion is concerned, however, none of the models presented acceptable discrimination (see Table 12 and Figure 5). Especially in Greece, the ROC curve area indicated a model that has negative discrimination, i.e., it is worse than random.

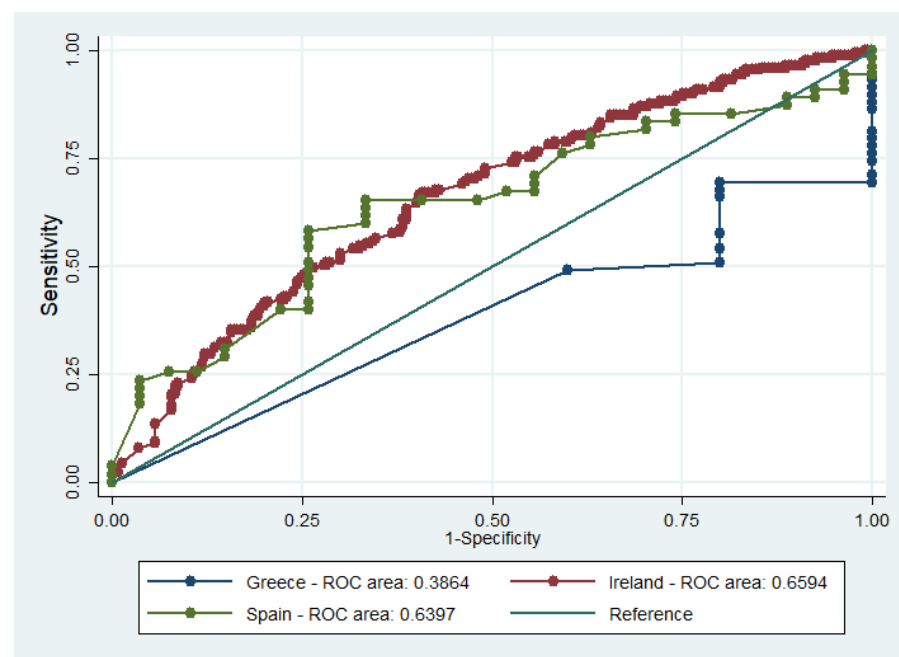
Table 11. Variables affecting firms' perceptions of energy efficiency—probit marginal effects (country subsamples).

Variables	Greece	Ireland	Spain
	Margins	Margins	Margins
Attitudes Towards EE			
Environmental impact reduction (=1 if strongly disagree)	-	0.056 *	0.026
		(0.029)	(0.084)
Chance on new technologies (=1 if strongly disagree)	-	0.038	0.124
		(0.032)	(0.083)
Lack access to loans, etc. (=1 if strongly disagree)	0.012	0.015	0.005
	(0.040)	(0.022)	(0.062)
Understand EE cost reduction (=1 if strongly disagree)	0.026	0.119 ***	-0.027
	(0.059)	(0.029)	(0.074)
Perception of Existing Label	-		
Energy labels affect my choices (=1 if strongly disagree)	-	0.024	-0.001
		(0.029)	(0.097)
Pseudo-R ²	0.012	0.057	0.027
Pearson's χ^2	0.156	0.659	0.536
Observations	47	409	82

Note: Standard errors in parentheses *** $p < 0.01$, * $p < 0.1$.

Table 12. Comparison of the ROC areas—country subsample models (firms).

Country	Obs.	ROC Area	Std. Err.	Asymptotic Normal (95% Conf. Interval)	
Greece	64	0.386	0.101	0.189	0.583
Ireland	409	0.659	0.027	0.607	0.712
Spain	82	0.640	0.063	0.516	0.764

**Figure 5.** ROC curves and C-statistic estimates for the national models (firms).

Since the national models failed to provide adequate discrimination, no further tests were carried out. All in all, it could be argued that the model is not transferable across countries. However, this conclusion should be seen with caution owing to the small number of observations in the cases of Spain and Greece.

3.4.2. Technology Transferability

The results for the three technological subsamples are given in Tables 13 and 14. The calibration criterion was generally satisfied for all countries; therefore, the null hypothesis that the logistic regression model used is the correct model could not be rejected at a 5% level. Based on Pearson's χ^2 test, it is suggested that there was no difference between the ROC areas of the three models ($p = 0.729$) (Table 15 and Figure 6). However, the predictive power of the appliances model was barely acceptable.

Table 13. Variables affecting firms' perceptions of energy efficiency—probit marginal effects (technology subsamples).

Variables	Appliances	Properties	Transport
	Margins	Margins	Margins
Attitudes Towards EE			
Environmental impact reduction (=1 if strongly disagree)	0.035 (0.063)	0.053 (0.043)	0.026 (0.084)
Chance on new technologies (=1 if strongly disagree)	0.115 * (0.061)	0.015 (0.048)	0.124 (0.083)
Lack access to loans, etc. (=1 if strongly disagree)	0.007 (0.034)	0.046 (0.034)	0.005 (0.062)
Understand EE cost reduction (=1 if strongly disagree)	0.060 (0.040)	0.010 (0.039)	−0.027 (0.074)
Perception of Existing Label			
Energy labels affect my choices (=1 if strongly disagree)	−0.039 (0.057)	0.083 * (0.049)	−0.001 (0.097)
Pseudo-R ²	0.074	0.064	0.048
Pearson's χ^2	0.176	0.403	0.517
Observations	146	144	265

Note: Standard errors in parentheses * $p < 0.1$.

Table 14. Comparison of the ROC areas—technology subsample models (firms).

Technology	Obs.	ROC Area	Std. Err.	Asymptotic Normal (95% Conf. Interval)	
Appliances	146	0.703	0.055	0.595	0.811
Properties	144	0.682	0.050	0.583	0.780
Transport	265	0.654	0.034	0.588	0.720

Table 15. Wald test results for the pooled and technological subsample models (firms).

Country Difference	Prob > χ^2
Pooled–Appliances	0.5020
Pooled–Properties	0.2330
Pooled–Transport	0.0057
Appliances–Properties	0.2614
Appliances–Transport	0.9300
Properties–Transport	0.1460

Although the discrimination of the models was not satisfactory, the “transferability” of the model was examined by employing the Wald test (Table 15). The null hypothesis was rejected between the pooled model and the transport model, just like in the household survey. Therefore, it could be argued that the model is transferable as regards the appliances and properties models, but not for the transport model. Yet, as mentioned, the models had weak predictive power.

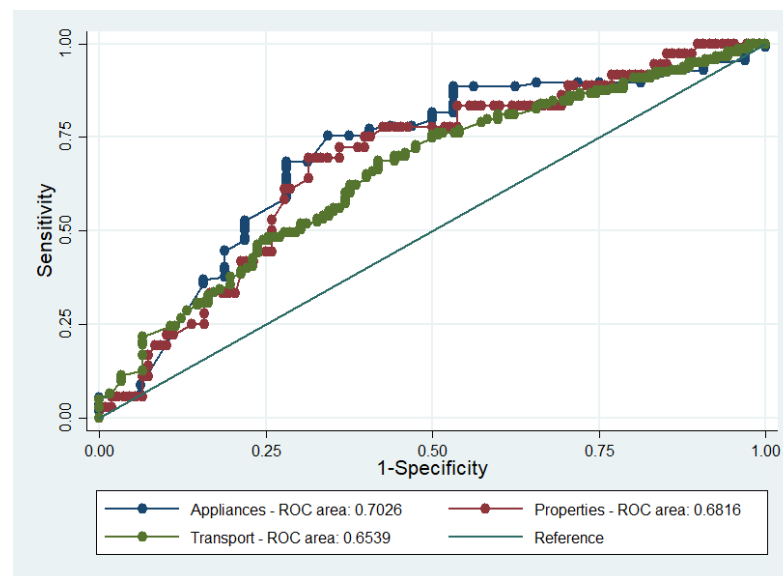


Figure 6. ROC curves and C-statistic estimates for the technological models (firms).

3.4.3. Sector Transferability

The “transferability” of the model was finally tested using the sectoral subsamples (i.e., agriculture and services). The results are presented in Tables 16 and 17. The calibration criterion was satisfied for both sectoral models, (i.e., the hypothesis that the correct model is used could not be rejected). Based on Pearson’s χ^2 test, the null hypothesis that there is no difference between the ROC areas of the two models was rejected ($p = 0.015$). Furthermore, only the services model showed an acceptable C-statistic value (Table 17 and Figure 7). The “transferability” of the model was examined by means of the Wald test (Table 18). The null hypothesis was rejected for all pair comparisons. Therefore, it can be concluded that the model is not transferable across sectors.

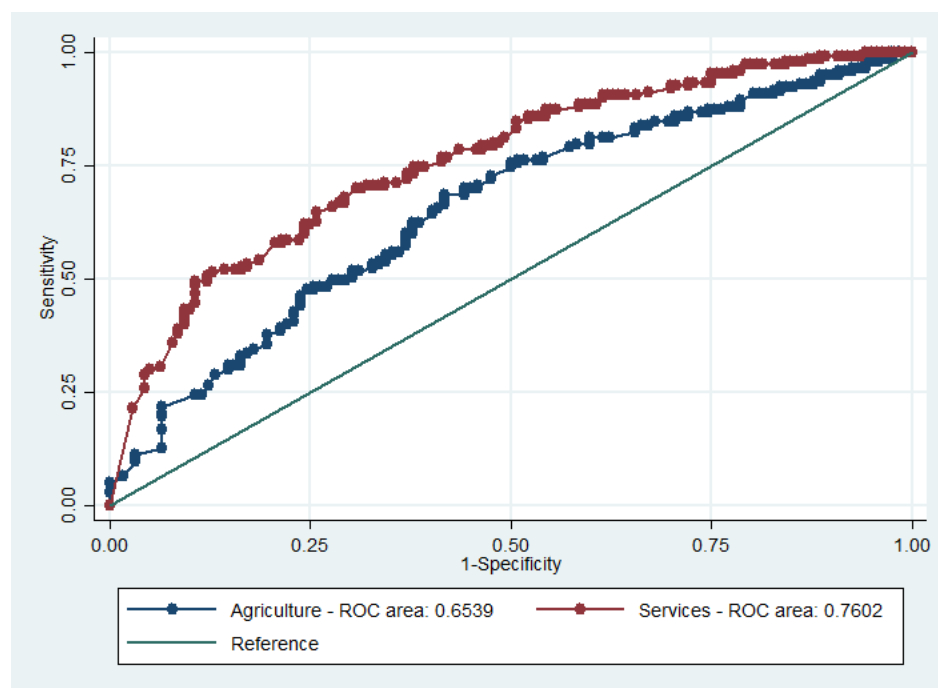
Table 16. Variables affecting firms’ perceptions of energy efficiency—probit marginal effects (sectoral subsample).

Variables	Agriculture	Services
	Margins	Margins
Attitudes Towards EE		
Environmental impact reduction (=1 if strongly disagree)	0.037 (0.037)	0.095 ** (0.044)
Chance on new technologies (=1 if strongly disagree)	0.066 (0.040)	0.096 ** (0.048)
Lack access to loans, etc. (=1 if strongly disagree)	−0.016 (0.027)	0.099 *** (0.028)
Understand EE cost reduction (=1 if strongly disagree)	0.118 *** (0.042)	0.075 ** (0.035)
Perception of Existing Label		
Energy labels affect my choices (=1 if strongly disagree)	−0.005 (0.036)	0.089 ** (0.043)
Pseudo-R ²	0.048	0.162
Pearson’s χ^2	0.517	0.453
Observations	265	290

Note: Standard errors in parentheses *** $p < 0.01$, ** $p < 0.05$.

Table 17. Comparison of the ROC areas—sectoral subsample models (firms).

Technology	Obs	ROC Area	Std. Err.	Asymptotic Normal (95% Conf. Interval)	
Agriculture	265	0.654	0.034	0.588	0.720
Services	290	0.760	0.028	0.706	0.815

**Figure 7.** ROC curves and C-statistic estimates for the sectoral models (firms).**Table 18.** Wald test results for the pooled and sectoral subsample models (firms).

Technology Difference	Prob > χ^2
Pooled–Agriculture	0.0057
Pooled–Services	0.0125
Agriculture–Services	0.0040

4. Conclusions

So far, several research efforts have been carried out to investigate the role of socioeconomic characteristics and behavioral and market biases on energy efficiency investment decisions. Despite this wealth of information, however, the transferability of the results between countries or sectors remains unexplored. This is an important shortcoming because the demand for evidence-based energy policymaking is high. Too often, though, the policymakers must rely on previous studies to obtain reliable and comparative data and formulate policies at a new site, since the cost of generating new evidence for policy sites is prohibited. Since the 2000's this is the task of the emerging field of policy transfer [60], which aims at approximating the policy variables at a new site (policy site) by using the estimated values at a different but similar area, region or country (study site) [61]. In all areas of policy transfer though, the transferability of policies between countries has received limited attention [39,62,63], and this stands also for the energy policy field [39,40].

Aiming to fill this gap and addressing the inherent difficulties in cross-country policy transfer, this study applied state-of-the-art econometric techniques to complement the mostly qualitative [60] methodological tools of policy transfer. To this end, the generaliz-

ability of the theoretical and econometric models developed in the context of the H2020 CONSEED project for investigating behavioral anomalies in EU energy efficiency was tested. The surveys were designed in a harmonized way to facilitate subsequent data pooling where possible. Specifically, two homogenized datasets for validation purposes and evaluation of the “transferability” of the models across different countries, technologies, and sectors were constructed. To the authors’ knowledge, this is the first attempt to examine whether the relationships between energy efficiency-related attitudes, sociodemographic characteristics, and potential market and non-market failures are general or attributable only to a narrow set of conditions. The novelty of our approach lies with (1) providing a robust statistical specification of heterogeneity in the consumer, technology, and sector space and (2) understanding in a quantitative, statistical sense the limitations of lesson drawing in a policy transfer exercise. The main conclusions drawn from the statistical process are discussed hereinafter.

According to the C-statistic—a measure of the predictive power of the model—and Pearson’s χ^2 goodness-of-fit statistic—a measure of calibration of the model—the regression model used was the correct one; the pooled models derived for the households and the firm samples passed the internal validation tests (that is, Proposition 1 was verified). Further, given the larger number of observations and the diverse nature of the population, the pooled samples seemed to allow for a more reliable investigation of the relative importance of the factors influencing consumers’ attitudes and beliefs towards energy investment decisions.

Our results show that energy efficiency importance varied across products and countries. For instance, in household appliances, energy efficiency was among the top two factors (more than 70% rated it as very important), whereas energy efficiency for cars was ranked only fourth. For this reason, and as expected from the relevant literature, the transferability of the results (i.e., the use of a “universal” model of energy efficiency from the pooled model), across countries, technologies, and sectors is limited. This was confirmed by the statistical tests used to determine transferability (i.e., Proposition 2 was rejected). Further, relatively small sample sizes for some sectors may also have contributed to low transferability.

All in all, the pooled models were validated, i.e., they were correctly specified and accurately represented the phenomena studied. Nevertheless, the extrapolation of the results to specific populations in terms of “space” (i.e., country) and “target” (e.g., sectors and technologies) should be seen with caution from a policy perspective for two reasons. First, as already mentioned, the transferability of the models proved to be limited owing to the plethora of parameters influencing energy efficiency-related attitudes and behavior. This finding is consistent with the results of previous research. For instance, Wang et al. [38], who compared seven data-driven models to forecast electricity usage at a city-scale level, concluded that the models can predict the city-level electricity demand well but their generalizability is in question. Miller [36] used machine learning modeling approaches in building energy prediction and found that no single modeling technique produces accurate results for all the different types of buildings. Hopf et al. [35] used machine learning models to investigate, among other things, whether a model trained on one geographic region can be transferred to another geographic region. The authors claimed that the model can be transferred “... with acceptable performance loss ...” Nevertheless, when they tested whether data from one company can improve the predictiveness of the model of another company, they noticed that “... no performance improvement was possible ...” Finally, Martinez-Soto and Jentsch [37], who developed and tested a transferable residential energy model, found that their model could predict the final energy demand of the residential sector of different countries. Yet, they noticed that the model relies only on six input parameters and, thus, it was not possible to check the effect of detailed energy efficiency measures that could be implemented in a country. Second, one cannot exclude the possibility of endogeneity, i.e., the correlation of an explanatory variable in the regression model with the error term in the model [64–66]. Two cases are especially

common in social research: (1) when important variables are omitted from the model (called “omitted variable bias”) and (2) when the dependent variable is also a predictor of an independent variable (known as “simultaneity bias” or “reverse causality”) [65]. A common method for dealing with endogeneity is the use of instrumental variables, but in cross-sectional surveys and with few variables, like in our case, it is practically impossible to handle endogeneity [65]. Similarly, controlling for all sources of variance in the context of social science is not feasible, i.e., it is not possible to include all possible sources of variance of a dependent variable in the regression model from a practical viewpoint [64]. For this reason, and even though all surveys were based on randomized samples, a design that assures the causal inference [64,67] endogeneity may still be an issue (i.e., those who believe that “energy efficiency is a very important attribute in the purchasing decision” may also be more affected by the energy labels, i.e., the two variables are simultaneously causing each other, and this may bias the results upwards as well as downwards). To provide more consistent answers, future research surveys need to collect different subsamples for each country and for several products that will allow international and intersectoral comparisons.

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/en14113122/s1>.

Author Contributions: Conceptualization: A.K., D.D., and M.S.; data curation: D.D., C.T., E.D., I.G., S.K., and E.L.; formal analysis: A.K., D.D., C.T., and M.S.; methodology: A.K., D.D., and M.S.; project administration: E.D. and M.S.; writing: A.K., D.D., C.T., and M.S.; review and editing: E.D., I.G., S.K., and E.L.; All authors have read and agreed to the published version of the manuscript. The manuscript solely reflects the authors’ views.

Funding: This work was supported by the CONSEED Project (CONsumer Energy Efficiency Decision making), funded under the HORIZON Framework Programme of the European Commission (contract no. 723741).

Institutional Review Board Statement: Ethical review and approval were waived for this study, because the surveys were conducted before the entry into force of the General Data Protection Regulation (Regulation No 2016/679).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available because ownership of background information remains with the originating partner.

Conflicts of Interest: The authors declare that they have no conflict of interest.

References

1. IEA. *World Energy Outlook 2020: Part of World Energy Outlook*; International Energy Agency: Paris, France, 2020.
2. OECD. *Building Back Better—A Sustainable, Resilient Recovery after COVID-19*; Organisation for Economic Cooperation and Development: Paris, France, 2020; p. 15.
3. Blasch, J.; Filippini, M.; Kumar, N. Boundedly rational consumers, energy and investment literacy, and the display of information on household appliances. *Resour. Energy Econ.* **2019**, *56*, 39–58. [[CrossRef](#)]
4. Newell, R.G.; Siikamäki, J. Individual Time Preferences and Energy Efficiency. *Am. Econ. Rev.* **2015**, *105*, 196–200. [[CrossRef](#)]
5. Poortinga, W.; Steg, L.; Vlek, C.; Wiersma, G. Household preferences for energy-saving measures: A conjoint analysis. *J. Econ. Psychol.* **2003**, *24*, 49–64. [[CrossRef](#)]
6. Solà, M.D.M.; De Ayala, A.; Galarraga, I.; Escapa, M. Promoting energy efficiency at household level: A literature review. *Energy Effic.* **2021**, *14*, 6. [[CrossRef](#)]
7. Gerarden, T.D.; Newell, R.G.; Stavins, R.N. Assessing the Energy-Efficiency Gap. *J. Econ. Lit.* **2017**, *55*, 1486–1525. [[CrossRef](#)]
8. Cattaneo, C. Internal and external barriers to energy efficiency: Which role for policy interventions? *Energy Effic.* **2019**, *12*, 1293–1311. [[CrossRef](#)]
9. Lillemo, S.C. Measuring the effect of procrastination and environmental awareness on households’ energy-saving behaviours: An empirical approach. *Energy Policy* **2014**, *66*, 249–256. [[CrossRef](#)]
10. Ouyang, J.; Gao, L.; Yan, Y.; Hokao, K.; Ge, J. Effects of Improved Consumer Behavior on Energy Conservation in the Urban Residential Sector of Hangzhou, China. *J. Asian Arch. Build. Eng.* **2009**, *8*, 243–249. [[CrossRef](#)]

11. Lesic, V.; De Bruin, W.B.; Davis, M.C.; Krishnamurti, T.; Azevedo, I.M.L. Consumers' perceptions of energy use and energy savings: A literature review. *Environ. Res. Lett.* **2018**, *13*, 033004. [[CrossRef](#)]
12. De Ayala, A.; Foudi, S.; Solà, M.D.M.; López-Bernabé, E.; Galarraga, I. Consumers' preferences regarding energy efficiency: A qualitative analysis based on the household and services sectors in Spain. *Energy Effic.* **2021**, *14*, 3. [[CrossRef](#)]
13. Grondys, K.; Androniceanu, A.; Dacko-Pikiewicz, Z. Energy Management in the Operation of Enterprises in the Light of the Applicable Provisions of the Energy Efficiency Directive (2012/27/EU). *Energies* **2020**, *13*, 4338. [[CrossRef](#)]
14. Vlachokostas, C. Smart buildings need smart consumers: The meet-in-the middle approach towards sustainable management of energy sources. *Int. J. Sustain. Energy* **2020**, *39*, 648–658. [[CrossRef](#)]
15. Le Quéré, C.; Jackson, R.B.; Jones, M.W.; Smith, A.J.P.; Abernethy, S.; Andrew, R.M.; De-Gol, A.J.; Willis, D.R.; Shan, Y.; Canadell, J.G.; et al. Temporary reduction in daily global CO₂ emissions during the COVID-19 forced confinement. *Nat. Clim. Chang.* **2020**, *10*, 647–653. [[CrossRef](#)]
16. Koroleva, K.; Melenhorst, M.; Novak, J.; Gonzalez, S.L.H.; Fraternali, P.; Rizzoli, A.E. Designing an integrated socio-technical behaviour change system for energy saving. *Energy Inform.* **2019**, *2*, 30. [[CrossRef](#)]
17. Schleich, J.; Gassmann, X.; Meissner, T.; Faure, C. A large-scale test of the effects of time discounting, risk aversion, loss aversion, and present bias on household adoption of energy-efficient technologies. *Energy Econ.* **2019**, *80*, 377–393. [[CrossRef](#)]
18. Stadelmann, M. Mind the gap? Critically reviewing the energy efficiency gap with empirical evidence. *Energy Res. Soc. Sci.* **2017**, *27*, 117–128. [[CrossRef](#)]
19. Bianco, V.; Righi, D.; Scarpa, F.; Tagliafico, L.A. Modeling energy consumption and efficiency measures in the Italian hotel sector. *Energy Build.* **2017**, *149*, 329–338. [[CrossRef](#)]
20. Damigos, D.; Kontogianni, A.; Tourkolias, C.; Skourtos, M. Behind the scenes: Why are energy efficient home appliances such a hard sell? *Resour. Conserv. Recycl.* **2020**, *158*, 104761. [[CrossRef](#)]
21. Moon, J.B.; DeWitt, T.H.; Errend, M.N.; Bruins, R.J.F.; Kentula, M.E.; Chamberlain, S.J.; Fennessy, M.S.; Naithani, K.J. Model application niche analysis: Assessing the transferability and generalizability of ecological models. *Ecosphere* **2017**, *8*, e01974. [[CrossRef](#)]
22. Wenger, S.J.; Olden, J. Assessing transferability of ecological models: An underappreciated aspect of statistical validation. *Methods Ecol. Evol.* **2012**, *3*, 260–267. [[CrossRef](#)]
23. Peterson, A.T.; Papeş, M.; Eaton, M. Transferability and model evaluation in ecological niche modeling: A comparison of GARP and Maxent. *Ecography* **2007**, *30*, 550–560. [[CrossRef](#)]
24. Patton, A.P.; Zamore, W.; Naumova, E.N.; Levy, J.I.; Brugge, D.; Durant, J.L. Transferability and Generalizability of Regression Models of Ultrafine Particles in Urban Neighborhoods in the Boston Area. *Environ. Sci. Technol.* **2015**, *49*, 6051–6060. [[CrossRef](#)]
25. Allen, R.W.; Amram, O.; Wheeler, A.; Brauer, M. The transferability of NO and NO₂ land use regression models between cities and pollutants. *Atmos. Environ.* **2011**, *45*, 369–378. [[CrossRef](#)]
26. Poplawski, K.; Gould, T.; Setton, E.; Allen, R.; Su, J.; Larson, T.; Henderson, S.; Brauer, M.; Hystad, P.; Lightowlers, C.; et al. Intercity transferability of land use regression models for estimating ambient concentrations of nitrogen dioxide. *J. Expo. Sci. Environ. Epidemiol.* **2008**, *19*, 107–117. [[CrossRef](#)]
27. Sikder, S.; Pinjari, A. Spatial Transferability of Person-Level Daily Activity Generation and Time Use Models: Empirical Assessment. *Transp. Res. Rec.* **2013**, *2343*, 95–104. [[CrossRef](#)]
28. Kawamoto, E. Transferability of standardized regression model applied to person-based trip generation. *Transp. Plan. Technol.* **2003**, *26*, 331–359. [[CrossRef](#)]
29. Sikder, S.; Pinjari, A.R.; Srinivasan, S.; Nowrouzian, R. Spatial transferability of travel forecasting models: A review and synthesis. *Int. J. Adv. Eng. Sci. Appl. Math.* **2013**, *5*, 104–128. [[CrossRef](#)]
30. Memon, M.A.; Matsuoka, S. Benefit Transfer Function to Estimate WTP for Rural Water Supply in Pakistan. *J. Int. Dev. Stud.* **2001**, *10*, 101–119. [[CrossRef](#)]
31. Rozan, A. Benefit Transfer: A Comparison of WTP for Air Quality between France and Germany. *Environ. Resour. Econ.* **2004**, *29*, 295–306. [[CrossRef](#)]
32. Downing, M.; Ozuna, J.T. Testing the Reliability of the Benefit Function Transfer Approach. *J. Environ. Econ. Manag.* **1996**, *30*, 316–322. [[CrossRef](#)]
33. Leon-Gonzalez, R.; Scarpa, R. Improving multi-site benefit functions via Bayesian model averaging: A new approach to benefit transfer. *J. Environ. Econ. Manag.* **2008**, *56*, 50–68. [[CrossRef](#)]
34. Ferrini, S.; Schaafsma, M.; Bateman, I. Revealed and stated preference valuation and transfer: A within-sample comparison of water quality improvement values. *Water Resour. Res.* **2014**, *50*, 4746–4759. [[CrossRef](#)]
35. Hopf, K.; Riechel, S.; Sodenkamp, M.A.; Staake, T. Predictive Customer Data Analytics—The Value of Public Statistical Data and the Geographic Model Transferability. In Proceedings of the ICIS 2017: Transforming Society with Digital Innovation, Seoul, Korea, 10–13 December 2017.
36. Miller, C. More Buildings Make More Generalizable Models—Benchmarking Prediction Methods on Open Electrical Meter Data. *Mach. Learn. Knowl. Extr.* **2019**, *1*, 974–993. [[CrossRef](#)]
37. Martinez-Soto, A.; Jentsch, M.F. A transferable energy model for determining the future energy demand and its uncertainty in a country's residential sector. *Build. Res. Inf.* **2019**, *48*, 587–612. [[CrossRef](#)]

38. Wang, Z.; Hong, T.; Li, H.; Piette, M.A. Predicting city-scale daily electricity consumption using data-driven models. *Adv. Appl. Energy* **2021**, *2*, 100025. [\[CrossRef\]](#)
39. Warren, P. Transferability of demand-side policies between countries. *Energy Policy* **2017**, *109*, 757–766. [\[CrossRef\]](#)
40. Bößner, S.; Suljada, T.; Johnson, F.X.; Bruno, A.; Morales, J.R.; Hu, M.; Bhamidipati, P.L.; Haselip, J. Policy transfer processes and renewable energy penetration: A comparative analysis of Peru, Thailand, and Uganda. *Sustain. Earth* **2020**, *3*, 2. [\[CrossRef\]](#)
41. Orlov, A.; Kallbekken, S. The impact of consumer attitudes towards energy efficiency on car choice: Survey results from Norway. *J. Clean. Prod.* **2019**, *214*, 816–822. [\[CrossRef\]](#)
42. Lakić, E.; Damigos, D.; Gubina, A.F. How important is energy efficiency for Slovenian households? A case of homeowners and potential homebuyers and their willingness to invest in more efficient heating controls. *Energy Effic.* **2021**, *14*, 9. [\[CrossRef\]](#)
43. López-Bernabé, E.; Foudi, S.; Linares, P.; Galarraga, I. Factors affecting energy-efficiency investment in the hotel industry: Survey results from Spain. *Energy Effic.* **2021**, *14*, 41. [\[CrossRef\]](#)
44. Foudi, S.; de Ayala, A.; López-Bernabé, E.; Galarraga, I. *Consumer Survey Based Empirical Evidence on Consumer's Energy Efficiency Choices across Different Consumer Groups and Geographical Locations: Survey Results*; CONSEED: Dublin, Ireland, 2018; p. 118.
45. Shen, J. Understanding the Determinants of Consumers' Willingness to Pay for Eco-Labeled Products: An Empirical Analysis of the China Environmental Label. *J. Serv. Sci. Manag.* **2012**, *5*, 87–94. [\[CrossRef\]](#)
46. Allcott, H. Consumers' Perceptions and Misperceptions of Energy Costs. *Am. Econ. Rev.* **2011**, *101*, 98–104. [\[CrossRef\]](#)
47. Davis, L.W.; Metcalf, G.E. Does Better Information Lead to Better Choices? Evidence from Energy-Efficiency Labels. *J. Assoc. Environ. Resour. Econ.* **2016**, *3*, 589–625. [\[CrossRef\]](#)
48. Gillingham, K.; Harding, M.; Rapson, D. Split Incentives in Residential Energy Consumption. *Energy J.* **2012**, *33*, 37–62. [\[CrossRef\]](#)
49. Larrick, R.P.; Soll, J.B. The MPG Illusion. *Science* **2008**, *320*, 1593–1594. [\[CrossRef\]](#)
50. Turrentine, T.S.; Kurani, K.S. Car buyers and fuel economy? *Energy Policy* **2007**, *35*, 1213–1223. [\[CrossRef\]](#)
51. Vasseur, V.; Marique, A.-F.; Udalov, V. A Conceptual Framework to Understand Households' Energy Consumption. *Energies* **2019**, *12*, 4250. [\[CrossRef\]](#)
52. John, O.P.; Benet-Martínez, V. Measurement: Reliability, construct validation, and scale construction. In *Handbook of Research Methods in Social and Personality Psychology*, 2nd ed.; Cambridge University Press: New York, NY, USA, 2014; pp. 473–503, ISBN 978-1-107-01177-9.
53. Hosmer, D.W.; Lemeshow, S. *Applied Logistic Regression*, 2nd ed.; John Wiley & Sons: Hoboken, NJ, USA, 2000; ISBN 0-471-72214-6.
54. Steyerberg, E.W.; Eijkemans, M.J.C.; Harrell, J.F.E.; Habbema, J.D.F. Prognostic Modeling with Logistic Regression Analysis: In Search of a Sensible Strategy in Small Data Sets. *Med. Decis. Mak.* **2001**, *21*, 45–56. [\[CrossRef\]](#)
55. Arboretti Giancristofaro, R.; Salmaso, L. Model Performance Analysis and Model Validation in Logistic Regression. *Statistica* **2007**, *63*, 375–396. [\[CrossRef\]](#)
56. Harrell, F.E.J.; Lee, K.L.; Mark, D.B. Multivariable Prognostic Models: Issues in Developing Models, Evaluating Assumptions and Adequacy, and Measuring and Reducing Errors. *Stat. Med.* **1996**, *15*, 361–387. [\[CrossRef\]](#)
57. Blöchliger, A. Validation of Default Probabilities. *J. Financ. Quant. Anal.* **2012**, *47*, 1089–1123. [\[CrossRef\]](#)
58. DeLong, E.R.; DeLong, D.M.; Clarke-Pearson, D.L. Comparing the Areas under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach. *Biometrics* **1988**, *44*, 837–845. [\[CrossRef\]](#) [\[PubMed\]](#)
59. Judge, G.; Griffiths, W.; Hill, C.; Lütkepohl, H.; Lee, T.-C. The Theory and Practice of Econometrics. In *Wiley Series in Probability and Statistics*, 2nd ed.; John Wiley & Sons: Hoboken, NJ, USA, 1985.
60. Benson, D.; Jordan, A. What have we Learned from Policy Transfer Research? Dolowitz and Marsh Revisited. *Polit. Stud. Rev.* **2011**, *9*, 366–378. [\[CrossRef\]](#)
61. Evans, M. New directions in the study of policy transfer. *Policy Stud.* **2009**, *30*, 237–241. [\[CrossRef\]](#)
62. Bulmer, S.; Padgett, S. Policy Transfer in the European Union: An Institutional Perspective. *Br. J. Polit. Sci.* **2004**, *35*, 103–126. [\[CrossRef\]](#)
63. Brouwer, R.; Martin-Ortega, J.; Dekker, T.; Sardonini, L.; Andreu, J.; Kontogianni, A.; Skourtos, M.; Raggi, M.; Viaggi, D.; Pulido-Velazquez, M.; et al. Improving value transfer through socio-economic adjustments in a multicountry choice experiment of water conservation alternatives. *Aust. J. Agric. Resour. Econ.* **2015**, *59*, 458–478. [\[CrossRef\]](#)
64. Antonakis, J.; Bendahan, S.; Jacquart, P.; Lalive, R. On making causal claims: A review and recommendations. *Leadersh. Q.* **2010**, *21*, 1086–1120. [\[CrossRef\]](#)
65. Lynch, S.M.; Brown, J.S. Chapter 8—Stratification and Inequality Over the Life Course. In *Handbook of Aging and the Social Sciences*, 7th ed.; Handbooks of Aging; Binstock, R.H., George, L.K., Eds.; Academic Press: San Diego, CA, USA, 2011; pp. 105–117, ISBN 978-0-12-380880-6.
66. Nichols, A. Causal Inference with Observational Data. *Stata J.* **2007**, *7*, 507–541. [\[CrossRef\]](#)
67. Shadish, W.R.; Cook, T.D.; Campbell, D.T. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*; Houghton, Mifflin and Company: Boston, MA, USA, 2002; p. 623, ISBN 0-395-61556-9.