The Role of Individual Preferences to Explain the Energy Performance Gap

Salomé BAKALOGLOU[1] and Dorothée CHARLIER[2]

Abstract

The aim of this research is to quantify the role of socio-economic and individual preferences on energy performance gap in the residential sector. First, using a two-steps discrete-continuous choice method, we study the probability to live in an energy-efficient dwelling in order to identify the weight of household's characteristics and individuals' preferences on energy consumption. In the second part, we identify the drivers of the energy performance gap using a quantile regression. Quantile regression analysis let us to tease out the effects of preferences on entire distribution on energy performance gap spectrum instead of focusing on the conditional average. The purpose is to explain extreme behaviors regarding energy use. As a result, this research shows that individual characteristics contribute to an important part of the energy performance gap. In such a context, some warnings to public authorities are provided regarding the fair assessment of energy savings benefits after thermal retrofits.

Keywords: Residential energy consumption; Households 'preferences; Energy Performance Gap; Discrete-continuous choice method; Quantile regression.

JEL CODES: Q41; D12; C26; C21

[1] Salomé BAKALOGLOU, salome.bakaloglou@cstb.fr, Centre Scientifique et Technique du Bâtiment, 84 Avenue Jean Jaurès, 77420 Champs-sur-Marne

[2] Dorothée CHARLIER, dorothee.charlier@univ-smb.f, IREGE, Université de Savoie, 74940 Annecy le Vieux Cedex

1. Introduction

Energy efficiency in the residential sector is a significant lever for meeting 2020 energy goals. During the past years, driven by European commitments, public policies aiming at reducing energy consumption of residential sector such as thermal regulations for residential construction (RT2000, RT2005, and RT2012[3]) or thermal retrofits policies have been numerous[4] in France. Despite the existence of these measures, at national level, energy consumption due to space heating in permanently occupied dwellings has decreased by nearly 12% between 1990 and 2013 (CEREN[5]), demonstrating that there is still a lot of work to do to achieve national energy goals (-38% in 2020[6]). In order to achieve these goals, refurbishment measures and social interventions to encourage more efficient use of energy are potential solutions (Lopes et al. 2012).

In such a context, the aim of this research is to go further in the understanding of individual preferences on energy consumption. We also want to identify the drivers of the energy performance gap (the ratio of the real energy consumption on the theoretical energy consumption) in order to identify extreme behaviors regarding energy use. This improvement in our understanding is essential to obtain a clearer picture of the issues regarding energy performance and to adapt future policies. Energy behaviors represent a significant potential to increase energy efficiency in the building sector. Although energy behavior is a drive of energy demand, potential energy savings due to behavior are usually neglected in the economic literature (Lopes et al. 2012). In this paper, we firstly analyze, using a two-steps discrete-continuous choice method, the influence of preferences on energy consumption variability taking into consideration that the household's decision is divided in two parts. In a first step, the household decides to live in a housing unit according to its theoretical energy performance, then, in a second step, he decides how much energy to consume. Then, we go further by exposing the role of these drivers to explain extreme behaviors of under and over consumption using a quantile regression. Quantile regression analysis let us to tease out the effects of preferences on entire distribution on energy performance gap spectrum instead of focusing on the conditional average.

[3] http://www.rt-batiment.fr/batiments-neufs/reglementation-thermique-2012/presentation.html
[4] http://www.planbatimentdurable.fr
[5] http://www.ceren.fr/
[6] Loi Grenelle 1 https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000020949548

Economists have brought to light that energy consumption in residential sector depends on technical data for 40% (Belaïd 2016) but also on socio-economic characteristics such as revenue, household's age, tenure status, etc. for about 33% of final energy consumption. The role of behavioral characteristics on energy consumption variability (Belaïd 2016; Belaïd and Garcia 2016; Cayla et al. 2011) has also been highlighted. Understanding the determinants of energy consumption has been a recurrent topic of research over the past few years and is an important issue regarding the estimation of potential effects of energy policies. Indeed, the success of a policy for reducing energy consumption mainly depends of occupants' behavior. A policy or a measure to improve thermal performance can become inefficient if occupants change their behavior increasing their thermal comfort. Recent research has put into light the existence of an energy performance gap (Galvin et Sunnika-Blank 2012) between predicted energy consumption based on energy audit and real energy consumption. This gap could be partly explained by rebound effect, i.e. the increase of energy consumption (mainly driven by a higher comfort demand) after an energy efficiency improvement. However, the quantitative understanding of the gap drivers has been neglected by previous literature because of the lack of appropriate data.

The recent survey *PHEBUS*[7] with complete thermal data allows us to better understand household's behaviors and their influence on energy consumption variability through the measure of the energy performance gap. Some authors have shown that there was not only an absolute performance gap between theoretical and real energy consumption but also an apparent relationship between them (Allibe 2012; Galvin 2015). By using this new database, we have a formal assessment of theoretical energy performance with limited heterogeneity regarding measure. Indeed, all the dwellings in the database have been audited by the same company with the same method. Moreover, the theoretical values of performance are merely an estimation of the actual consumption, since they are based on standard values and do not take account of the lifestyle of the occupants. Thus, using as a reference the theoretical energy consumption, we can measure the "intensity of energy consumption[8]", i.e. the ratio between real consumption and theoretical consumption in order to explain extreme behaviors of over and under consumption. Understanding the energy performance gap is the core of this research and could be essential for policy makers. Information about households' preferences could also be used

[7] http://www.statistiques.developpement-durable.gouv.fr/sources-methodes/enquete-nomenclature/1541/0/enquete-performance-lhabitat-equipements-besoins-usages.html
[8] (Wirl 1987; Cayla, Maizi, et Marchand 2011)

to better inform energy models in order to provide more realistic predictions of the space heating demand and the potential energy savings from energy saving renovations.

The paper is organized as follows. Section 2 presents the literature review. Section 3 describes the data. The model and the results are respectively presented in section 4 and 5. The aims are the followings: (i) to better understand energy consumption determinants by taking into account interactions between theoretical energy performance and other determinants of energy consumption by using a discrete-continuous model and (ii) to explain extreme behaviors (over and restricted consumption) regarding energy use with quantile regressions. Finally, answering this two questions will allow us to make some policies recommendations for the future. Section 6 concludes.

2. Literature review

Final energy efficiency of a dwelling can be explained by three main determinants: technical building characteristics, households' characteristics and external factors (energy price, building environment, etc.). The role of individual preferences for comfort on energy consumption has also been highlighted but research on this topic is rarer. Here, we review effects of non-technical determinants on energy consumption.

Concerning the methodology, research on energy consumption determinants is divided in two parts: some study only direct effects of buildings and household's characteristics on energy consumption; others include indirect effects in their analyses by trying to capture interactions between thermal characteristics of the dwelling and households characteristics (discrete continuous model or others). These models assume that characteristics of households could play a twofold role on final energy consumption: firstly, via the choice of thermal characteristics of the home (buildings characteristics or appliances number). Indeed, household's individual characteristics could have an impact on the choice of thermal characteristics, which themselves directly influence final energy consumption (indirect effect). Then, household individual characteristics also play a direct role on final energy consumption variability regardless of thermal characteristics (direct effect). So, driving such discrete-continuous model allows authors to distinguish direct from indirect effects of determinants.

2.1 Determinants of energy consumption: global effects

In order to include relevant variables in our research, socio-economic and environmental determinants exposed in economic literature are reviewed and their effects on energy consumption. Main results are summarized in the Table 1 below.

Table 1: Drivers of energy consumption

Variables	Effects on energy consumption	Sources
Tenure status	Contrary to the theory which specifies that tenants are likely to consume more energy than owners (misaligned incentives), empirical works fails to find a consensus on the effect of tenure status on energy consumption.	(Belaïd 2016; Charlier 2015; Jones et al. 2015; Yohanis 2012)
Income	Income elasticity is positive in most of the studies, which is consistent with the "normal good status" of energy consumption; income elasticity stays low, often less than 0.15. But, the effect of household income is sometimes more complex. Although the poor use less energy, they have a relatively smaller opportunity of changing their equipment. A positive elasticity may involve mainly the purchase of more efficient energy equipments which will induce lower energy consumption.	(Cayla et al. 2011; Nesbakken 2001) (Belaïd 2016; Labandeira et al. 2006; Santamouris et al. 2007).
Number of occupants	Positive effect on energy consumption	(Leahy and Lyons 2010; Vaage 2000)
Age of the reference person	Age of the reference person a cycle effect on energy consumption correlated with the family's life cycle. Energy consumption is comparatively higher for dwellings between 45 and 65 than for the other age classes	(Belaïd 2016; Brounen and Kok 2011; Brounen et al. 2013) (Belaïd 2016)
Employment status	High employed professional consume less than others	(McLoughlin et al. 2012)
Rebound effect (due to energy efficiency improvement)	Improving energy efficiency of a dwelling (thanks thermal retrofits or replacement of non-efficient equipment) conducts to a relative increase of energy consumption. Additional energy consumption can reach 30%. Rebound effect can decrease real energy savings of thermal retrofits.	(A. Greening et al. 2000; Berkhout et al. 2000; Bourrelle 2014; Galvin 2015; Majcen et al. 2013; Sorrell and Dimitropoulos 2008; Thomas and Azevedo 2013b, 2013a) (Galvin 2014b)
Urban area type	No clear effect of housing density	(Belaïd 2016; Kaza 2010)

		(Belaïd 2016)
Energy price	Positive price-elasticity but estimates varies widely from -0.20 to -1.6	(Belaïd 2016; Dubin and McFadden 1984; B. Halvorsen and Larsen 2001; Labandeira et al. 2006; Larsen and Nesbakken 2004; Nesbakken 1999, 2001; Risch and Salmon 2017)
Preferences for comfort	Energy-saving behaviors and preferences for savings are drivers of energy efficiency. Savings potentials range from 1.1% to over 29%.	(Estiri 2015; Jones et al. 2016; Lopes et al. 2012)

2.2.1 Indirect determinants: a focus

Some authors assume that we have to take into account interactions between households and dwellings characteristics when assessing variability of energy demand (Dubin and McFadden 1984). Indeed, energy demand can be considered as an indirect product of household choice and consumption behavior. For instance, the household firstly chooses some specific dwellings' characteristics according to its needs that may influence then the energy level of the home (Estiri 2014, 2015). In 2006, Kriström (2006) explained that households do not demand energy "per se" but demand is combined with other goods such as "capital goods" (housing units). So, there exists two levels of analysis when we want to understand energy consumption determinants. Some authors exposed that households with higher income are more likely to live in big homes that consume more (Ewing and Rong 2008). More recently, Estiri (2015) spotlighted main interactions between buildings characteristics, lifecycle and socio-economic characteristics of the household and quantified direct and indirect roles of each in energy consumption with a covariance structure analysis. He reached the conclusion that main effects of socio-economic and lifecycle characteristics are carried out via building characteristics (expressed with a latent variable that includes surface, number of rooms and tenure status). Here, we make the assumption that household characteristics and energy efficiency level of home interact. We assume that household allocation to an energy efficient dwelling is more likely to happen with some individual characteristics of the household and with some specific living conditions "chosen" by the household (dwelling characteristics and environment). To test the assumption and extract direct contributions of individual characteristics on energy consumption variability, we use a two-step discrete continuous model with an ordered logit as first step. Then we include predicted probabilities of belonging to a specific energy class as explanatory variables of energy consumption. We want to see which part of energy consumption is due to households' preferences for comfort for space heating.

Although researches on determinants of final energy consumption are numerous, research on extreme behaviors regarding energy consumption are few (Lopes et al. 2012). Energy efficiency gap, i.e. differences between theoretical and real final energy consumption, is well described but not econometrically exploited. A such analysis lets us to identify who consume more than they should and who restrict their energy consumption to enjoy energy-savings.

2.2.2 The Energy performance gap

The energy performance gap is defined in the literature as the difference between intrinsic energy consumption assessed during energy audit and real energy consumption evaluated via energy bills (Allibe 2012; Galvin 2010; Galvin and Sunikka-Blank 2013; Galvin 2014a). This gap has been well studied since the ten last years but no many work has been done to highlight the precise role of the determinants quoted as potential influencing factors. Galvin (2012) did a primary descriptive work and a European comparison. After a review to expose international results on the topic, the authors focus on the case study of German dwellings and puts forward policy implications. As a result, also quoted by others (Berkhout et al. 2000; Majcen et al. 2013), a systematic trend is highlighted when theoretical measure of heating energy consumption (Energy Performance Certificate measurement) and the real one (final consumption based on bills) are compared. In one hand, the more performing the dwelling is, the more the gap between the two measures grows in the sense of over-consumption. This trend is explained as the consequence of the rebound effect: it is assumed that energy efficient dwellings have been renovated, leading to a behavioral change of occupants to adapt (i.e. increase) their heating comfort after thermal retrofits. In the other hand, the concept of "prebound effect" for the less energy efficient dwellings of which real heating energy consumption is systematically lower than the theoretical one (on average 30% lower) is introduced by Galvin (2012). Assumptions which have been made in the literature to explain the gap are the followings:

- Uncertainties in the calculation method itself (Allibe 2012; Galvin and Sunikka-Blank 2013, 2014) (Galvin et Sunnika-Blank 2012; Allibe 2012). It could be linked to errors in calculation or thermal model or incorrect assumptions (occupancy, technical factors…)

- Uncertainties if we compare EPC for two similar dwellings. They can come from human errors and subjectivity (when assessing quality or quantity of buildings materials, surfaces…). According to experts, 20 to 30% of uncertainty exists for the French EPC (Carassus et al. 2013). This kind of uncertainties will be reduced in our study because EPCs have been realized by the same certification agency.
- False assessment of real quality of energy installations because of non-observability. Indeed, to stay economically affordable, EPC energy audits are not probably deep enough to assess the real quality of dwellings. This explanation comes from technical studies (Carassus et al. 2013) or economics research (Allibe 2012).
- Influence of socio-economic and behavioral factors such as occupancy status, income level, number of occupants (which differs from EPC calculation assumption), or preferences (Cayla et al. 2011).

In this research, we will focus on the last assumption: we aim to identify individual characteristics and external factors influencing energy consumption with a focus on extreme behaviors.

3. Data and descriptive statistics

3.1 Data

To study the determinants of the gap between theoretical performance and real energy consumption, this research uses data coming from the "*PHEBUS*" survey, a national household energy survey conducted by the Department of Observations and Statistics (SOeS) depending on the French Ministry of Ecology and Sustainable Development. The survey contains over 2000 dwelling energy audits performed by the same company and driven in 2012, theoretical energy performance measures, real energy consumptions (based on energy bills) and social, economic and behavioral data of dwellings occupants. Data sets available through this survey as quite innovative as it allows us to get uniform assessments of Energy Performance Certificates (EPC) between dwellings.

In this dataset, information is also available on households' preferences. For each end-use (heating, hot-water and electricity), it is possible to know whether households favor comfort or energy savings. It is therefore possible to have a scale of preferences. A strong preference for comfort will be measured as a declared preference for each end-uses, a medium preference as a declared preference for two out of three end-uses and finally a low preference as a single declared preference for comfort. In this dataset, other variables can also be used as a proxy for comfort, for example, the heating temperature.

In the PHEBUS database, information is provided on the type of energy tariff (for gas and electricity) and the power required per type of fuel used (electricity, gas, oil). The power required and the type of energy tariff depend on the type of fuel used for the heating system and in consequences the energy mix as well as the number of rooms (or the surface area). Thus, it is possible to have different energy tariffs per energy mix's composition and the end-use of each energy among households.

But, no information is provided on the energy tariff. So, it is possible to improve the PHEBUS database with information on the energy cost and the subscription cost for each type of energy (oil, gas, electricity and wood) per the power required and the type of tariff in 2011 and 2012. This information (PEGASE database) is provided by the French Ministry of Energy and presented in appendix A (Table 6).

Finally, it is possible to calculate for each household, a weighted energy cost depending on the energy mix and the structure of energy consumption. With a weighted energy cost, we have a specific cost of energy for each household. The formulae is the following:

$$energy\ price_i = \sum_{j=1}^{n} \frac{volume\ in\ kwh_{ijt} \times energy\ price_{jt}}{total\ volume\ in\ kwh_i}$$

Where j represents the type of fuel, i the household and t the type of tariff for a specific energy (electricity or gas).

3.2 Descriptive statistics

Main descriptive statistics of variables used in the model are summarized in appendix B1 (Table 7).

3.2.1 Measure of theoretical and real energy performance of dwellings

Theoretical energy measure available in the "*PHEBUS*" survey is the Energy Performance Certificate (EPC). EPC certification includes an energy audit, realized by an approved auditor based on visual inspection and collection on technical data. This measure considers three energy uses: heating, hot water production and cooling. Nor lighting consumption nor domestic appliances are considered. Characteristics such as house construction data, windows and walls insulation, heating system performance and climate data are collected and merged to obtain an aggregated measure of energy consumption. Standardized occupant's energy behaviors are also integrated. EPC's result is a quantitative assessment of primary and primary energy consumption of the dwelling in kilowatt-hours. It also gives the ranking of dwellings in energy classes (seven classes, from A to G, available in appendix B2 Table 8 and Figure 5). One of the advantages of using EPCs values given in the "PHEBUS" database is that all dwellings energy audits are made by the same firm, mobilizing the same calculation method, which allows us in the end to get kind of "uniform" data. For this research, we use EPC measurement expressed in final energy to better match with real energy consumption. Measurement of real energy consumption is based on energy bills for the year 2012. Thus, it is possible to calculate the "intensity of energy use". This indicator lets us to identify over-consuming and under-consuming behaviors.

The "intensity of energy use (IEU)" indicator firstly described by Wirl (1987) and Cayla et al. (2011), is defined as follows:

$$IEU = \frac{Final\ energy\ consumption\ (based\ on\ energy\ bills)}{Theoretical\ energy\ consumption\ (EPC)}$$

When the ratio becomes bigger than one, it means that the dwelling is "over-consuming" in comparison with theoretical measure; if it is smaller, the dwelling is supposed to "under-consumed". The energy performance gap gathers the states where IEU significantly differs from one. This research aims to explain high and low measures of the ratio thanks to socio-economic and environmental data.

Results obtained from the IEU are summarized in Table 2 below.

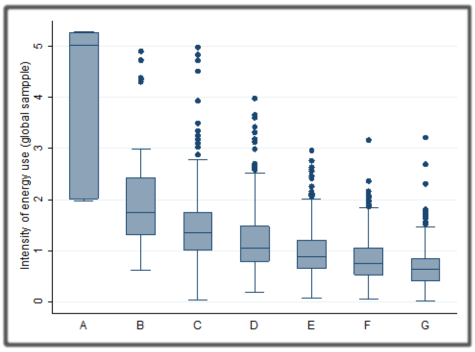
Table 2: Quantiles of intensity of energy use

	Mean	Std. Dev	Percentiles							Obs.
			5%	10%	25%	50%	75%	90%	95%	

IEU	1.05	0.629	0.292	0.44	0.649	0.925	1.312	1.765	2.117	2044

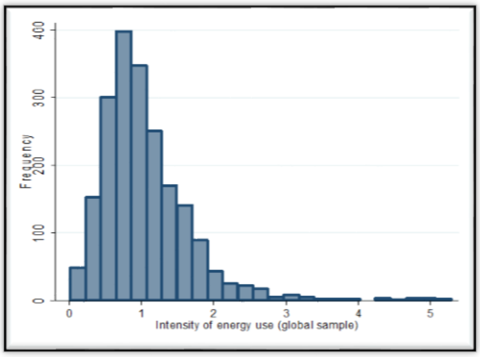
Analysis of "PHEBUS" data shows that intensity of energy use follows a visible trend linked to energy efficiency level of homes. For the less energy, efficient dwellings, energy is under-consumed, meaning that either theoretical energy measure is over assessed either households strictly restricts their energy consumption (Figure 1). The inverse trend is observable for very energy efficient dwellings: energy is over-consumed in energy classes A and B. It could mean that households living in energy efficient dwellings adapt their comfort and consumption behavior to the energy efficiency of their homes. This observation is usually called "rebound effect". Descriptive statistics of households characteristics (see Table 3 of descriptive statistics by energy class) make us set two assumptions: 1/ the fact of over-consuming could be linked to socio-demographic and economic characteristics or individual preferences for comfort of the households; 2/ the fact of living in an energy efficient dwelling could be determined by environmental constraints and socio-economic considerations (urban area, tenure status, availability on the market at the moving in time, price of the estate, household revenue). In our research, we aim to distinguish these effects to isolate direct effects of individual determinants on energy consumption.

Figure 1: Intensity of energy use



Data: PHEBUS 2012, authors calculus

Figure 2: Distribution of IEUs for the global sample



Data: PHEBUS 2012, authors calculus

Table 3: Descriptive statistics by energy class

Energy class	**A**	**B**	**C**	**D**	**E**	**F**	**G**
Number of observations	***5***	***42***	***281***	***563***	***600***	***298***	***248***
Average final energy consumption (kWh/m2/year) (based on bills)	177 (37.2)	125 (11.5)	137 (4.3)	148 (4.0)	173 (4.8)	196 (7.3)	218 (10.7)
Average EPC expressed in final energy	59 (19)	73 (3.7)	108 (1.9)	147 (2.8)	197 (3.9)	265 (7.4)	355 (12.7)
Average EPC expressed in primary energy	36 (1.9)	76 (1.4)	125 (1.2)	190 (1.2)	279 (1.3)	383 (2.3)	593 (10.5)
Intensity of energy use (global sample)	4 (0.7)	1.9 (0.17)	1.45 (0.05)	1.16 (0.03)	0.97 (0.03)	0.82 (0.03)	0.68 (0.03)
Intensity of energy use (gas sample)	0.97	1.64 (0.2)	1.03 (0.05)	0.79 (0.04)	0.69 (0.03)	0.59 (0.04)	0.52 (0.067)
Average annual disposable Income Per Household	53345 (9006)	46779 (7041)	42879 (1939)	40119 (1100)	35135 (888)	35041 (1803)	27489 (1146)
Average number of meter square by consumption unit (m2/CU)	100 (17.2)	81 (7.0)	71 (2.2)	69 (1.8)	67 (1.8)	61 (2.2)	57 (2.3)
Percentage of rented-occupied dwellings (%)	4	28	30	32	35	40	51
Mean heat temperature (°C)	20.4 (0.6)	20.2 (0.2)	20 (0.09)	20.4 (0.3)	21.2 (0.7)	21 (0.7)	20.7 (0.8)
Mean surface (m2)	182 (23.7)	142 (17.9)	120 (3.5)	110 (2.5)	99 (2.5)	88 (2.8)	76 (3.17)
Percentage of collective dwellings (%)	0	29	30	31	26	30	41
Average age of the reference person of the dwelling	56 (4.3)	52 (2.4)	50 (1.1)	54 (0.9)	56 (0.9)	56 (1.3)	57 (1.7)
Number of years spent in the current dwelling	9.6 (2.6)	9.6 (1.6)	12.7 (0.9)	14.7 (0.7)	16.9 (0.8)	17.8 (1.1)	17.5 (1.3)
Average number of appliance goods	18.8 (2.0)	19.5 (2.7)	15.9 (0.6)	16.1 (0.6)	15.2 (0.6)	13.5 (0.5)	11.9 (0.4)

Source: Author calculus, PHEBUS, 2012

3.2.2 Rebound and Prebound effects and link with individual preferences for comfort

The “PHEBUS” survey allows to go deeper in the understanding of the energy performance gap, especially via the consideration of the role played by individual preferences for comfort. Descriptive statistics have shown earlier that there was a systematic relationship between EPC measurement and real energy consumption: the more energy efficient the dwelling is, the higher the intensity of energy use is. That-is-to-say, over-consumption behavior is more likely to be found in the better efficient dwellings. We assume then that this trend could come from a “comfort effect”, i.e. an internalization and a behavioral adaptation of the fact of living in an efficient dwelling by households. In previous research, the observation has been linked to “rebound” and “prebound” effects (see literature review) and so associated with more important preferences for comfort of households. These preferences have been theoretically integrated into EPC calculation thanks to a factor of intermittence[9]: a better insulation implies an increase of theoretical energy consumption. Thanks to the precedent graphics (Figure 1), we can assume that this factor of intermittence is potentially under-estimated (link between IEU and EPC observable and negative).

The PHEBUS survey allows us to bring some new qualitative elements to characterize the link between individual preferences for comfort, higher energy consumption and theoretically energy performance of the dwelling. Using the variable asking household preferences for comfort over economy regarding three kinds of energy uses, we see that preferences for comfort are more likely to be found in more energy efficient dwellings (see Table 4 below), which is consistent with earlier stated assumptions and Figure 1. Moreover, a higher IEU seems correlated with a higher indoor heating temperature (Appendix B3 Table 9), which seems to demonstrate a relation between over-consumption and comfort preferences.

[9] It represents the variability of energy consumption due to day occupancy duration and also integrates a kind of “rebound effect”.

Table 3: Individual preferences for comfort over economy.

Percentage of households preferring comfort over economy for:	**A**	**B**	**C**	**D**	**E**	**F**	**G**
heating	85%	62%	54%	52%	56%	48%	54%
hot water	73%	70%	55%	57%	54%	50%	51%
specific electricity	27%	43%	46%	44%	40%	36%	39%
High preference for comfort*	21%	36%	31%	30%	30%	26%	26%
Medium preference for comfort*	58%	24%	19%	21%	19%	18%	19%
Low preference for comfort*	0%	18%	22%	20%	20%	21%	26%

**This variable is compounded from Phebus data: high preference for comfort means that household declared that it prefers comfort over economy for the three energy uses: specific electricity, heating and hot water, medium preference means that this preference for comfort concerns two of the three energy uses and finally, low preference means that the preference for comfort concerns only one energy use.*

Source: Phebus Survey 2012

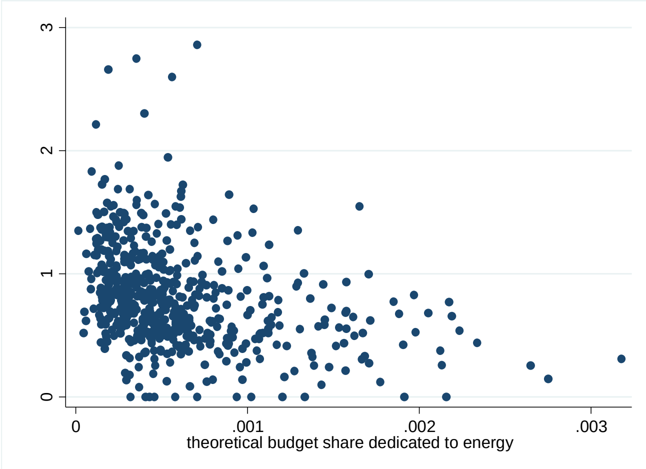
However, theoretical energy performance calculation can be biased because of measurement and could also partly constitute a potential explanation for these significant results. Moreover, as trend in IEU also exists in sample without having realized energy retrofits, we exclude the fact that it is totally explained by rebound effect due to retrofit achievements.

Finally, we try to link the energy performance gap to other determinants and use the indicator called "theoretical budget share dedicated to energy" introduced by Allibe (2012). This indicator is defined as:

$$EBS = \frac{Theoretical\ annual\ energy\ consumption\ (EPC) * energy\ price}{household\ income}$$

We observe that intensity of energy use (Figure 3) decreases with a higher theoretical budget share. According to the definition of this indicator, it means that people with higher revenue or living in more energy efficient dwelling (or both) increase their energy consumption in comparison to households with low revenue or non-efficient dwelling. As a conclusion of this descriptive analysis, we can set the following assumption: "comfort effect" and socio-economic determinants are potential drivers of the energy performance gap.

Figure 3: Evolution of intensity of energy use with theoretical budget share dedicated to energy



4. Model

In this paper, we proceed in two steps:

- We firstly analyze, using a two-steps discrete-continuous choice method, the influence of preferences on energy consumption variability taking into consideration that the household's decision is divided in two parts. In a first step, the household decides to live in a housing unit according to its theoretical energy performance, then, in a second step, he decides how much energy to consume.
- Then, we go further by exposing the role of these drivers to explain extreme behaviors of under and over consumption using a quantile regression. Quantile regression analysis let us to tease out the effects of preferences on entire distribution on energy performance gap spectrum instead of focusing on the conditional average.

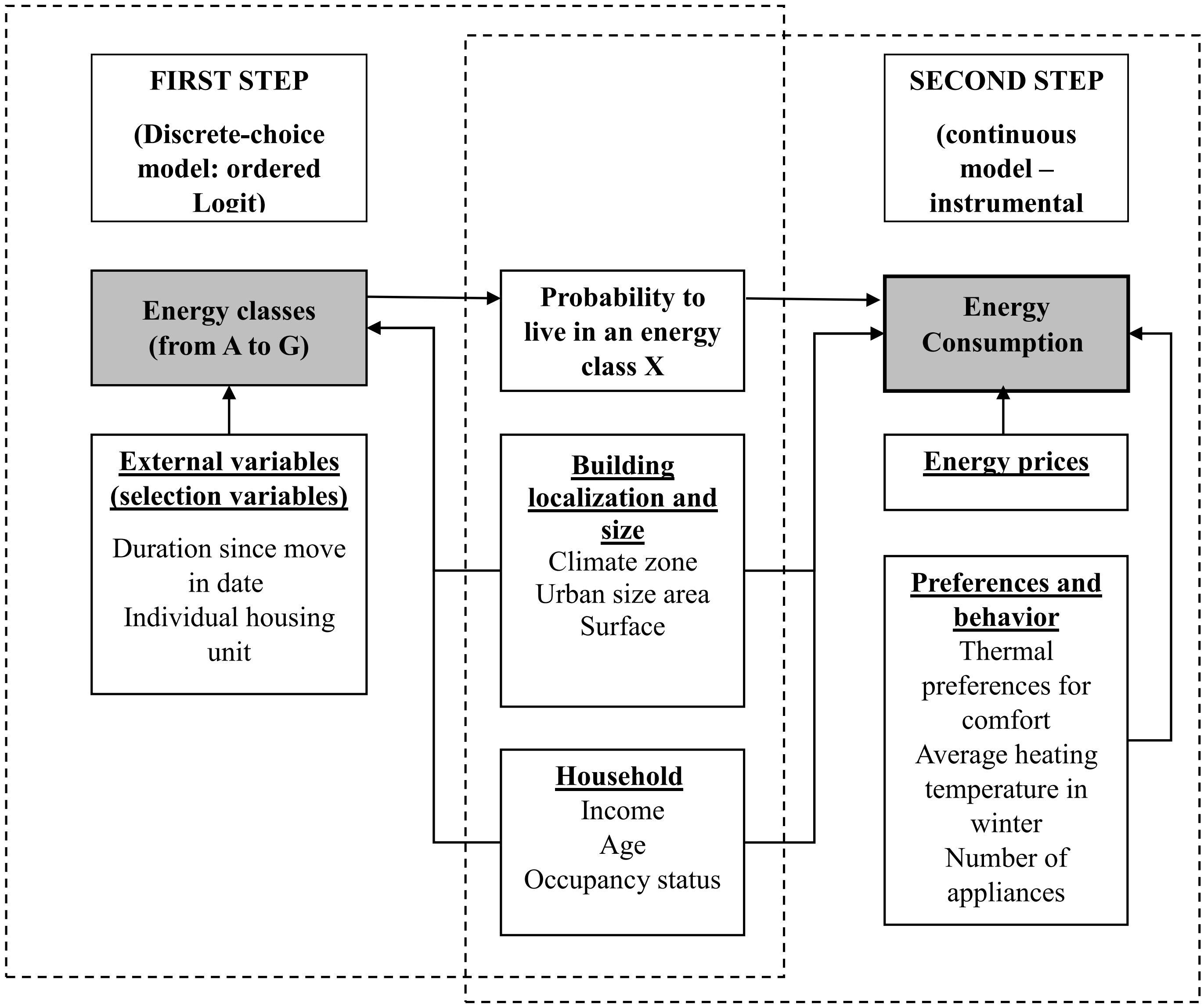
4.1 Two-stage model

Bottom-up economic models used to compute residential energy consumption are twofold. On one hand, some of the authors choose to study directly energy consumption by regressing it on its potential determinants using standard least squares estimation (Leahy and Lyons 2010; Rehdanz 2007). Sometimes, households or dwellings classification are computed as first step of the model and interclass-variability is analyzed (Belaïd and Garcia 2016). Others, by using a conditional demand analysis, assess energy consumption conditionally to the appliance stock which are in place (Aydinalp et al. 2002). On the other hand, discrete-continuous two steps models can be used. In these sequential models, authors consider that appliance or thermal equipment choices and consumption choice are bound and need to be linked via econometric assessment (Dubin and McFadden 1984; Risch and Salmon 2017; Vaage 2000). Here, in a similar vein, we make the assumption that theoretical energy performance level interacts with socio-economic and external determinants of the dwelling/household which themselves also directly influence final energy consumption.

Discrete-continuous methods are used to address selectivity biases in data sets with endogenously partitioned observational units (Frondel et al. 2016). These two-steps models are often used in the field of energy consumption because of interactions and endogeneity between independent explanatory variables. These two-step methods firstly consist in employing a multinomial logit to obtain selection terms which are then used as explanatory variables in the second stage. In the field of energy consumption in the residential sector, some research has been done on one specific kind of model: using heating equipment choice as discrete step, authors then study energy consumption determinants as second step (Dubin and McFadden 1984; Risch and Salmon 2017). In this research, an original data sets is used to apply this two-step method. As choice variable for the discrete step, we propose to use theoretical energy performance of the dwelling in classes. Energy classes are ordered and this specific element is considered in the econometric analysis.

By using a two-steps discrete-continuous method, we assume that efficiency level of the home and individual characteristics of the household interact. Moreover, in this paper, we focus on a discrete-continuous model because we face one potential problems of endogeneity: this is related to the choice of thermal performance of the dwelling (energy classes). Firstly, we study what determines the allocation of a dwelling/household to an energy efficient dwelling and we estimate for each observation a probability of living in such a dwelling. Then, we integrate this result in the second step of the analysis to determine if this probability impacts energy consumption to capture assumed interactions and extract the real weights of each determinants in global energy consumption. Moreover, we take into account a second potential endogeneity problem related to energy prices. Therefore, we introduced as explanatory variable in the second step the log of energy prices. To this end, we use instrumental variables to estimate the energy consumption choice, and we choose as instruments the energy prices in 2011 and the type of energy tariff for electricity. This last step let us to determine what is due to thermal performance and what is due to preferences and behavior. The discrete-continuous decision method makes it possible to account for the interrelationship between the choice of energy label cateogry and the quantity of energy use. This two-stage model is largely used to correct for endogeneity of discrete variables (Heckman and Robb 1985). The model is presented in Figure 4.

Figure 4: formalization of the econometric model



4.1.1 The Discrete Choice

The aim of this first model is to explain final energy consumption by considering that the fact of living in an energy efficient dwelling can interact with other factors such as household characteristics, dwelling environment and individual preferences. Energy class of the dwelling coming from EPC assessment is chosen as proxy for the theoretical energy efficiency level of the dwelling. Our model contains variables which are supposed to explain both choices: the choice of a dwelling with a certain energy efficiency level and the choice of energy use. For the first stage of the model, we use an ordered logit because energy performance class arises sequentially (Cameron and Trivadi 2010). For individual i, we specify:

$$y_i^* = x_i'\beta + u_i$$

With y^*a latent variable which is an unobserved measure of energy performance of the dwelling; x the regressors. For low y^*, energy performance is very high; for $y^* > \alpha_1$ corresponding to the energy class's threshold A-B to C, energy performance is a little bit lower; for $y^* > \alpha_2$ corresponding to the change C to D energy efficiency is even lower, etc. For a m-alternative ordered model (here $m = 6$ because of the 6 energy classes we consider), we define:

$$y_i = j \quad if \;\; \alpha_{j-1} < y_i^* \leq \alpha_j, \qquad j = 1, \dots, m$$

$$\Pr(y_i = j) = \Pr(\alpha_{j-1} < y_i^* \leq \alpha_j)$$

The regression parameters β and the m-1 threshold parameters $\alpha_1, \dots, \alpha_{m-1}$ are obtained by maximizing the log likelihood with $p_{ij} = \Pr(y_i = j)$. Then, we can compute probabilities of choosing the alternative j considering the regressors x. Predict probabilities are then integrated to the second step and used as regressors of final energy consumption expressed in kW/m^2/year with other explanatory variables. The model captures the possibility of correlation between unobservable variables in the discrete and the continuous stages.

4.1.2 *The continuous choice*

Conditional on this previous choice, a household decides the quantity of energy to consume. Therefore, in the second stage, the total energy consumption (the logarithm of the energy consumption in kWh/m^2) is estimated, conditional on the dwelling thermal performance (energy class). This is the " energy consumption choice." The joint estimation of both choices enables us to capture the potential correlation between unobservable variables in the discrete and the continuous stages. We estimate it using a least square model. To control endogeneity of the energy prices variable in 2012 (P_{2012i}), we introduce as instruments the lag of energy prices (P_{2011i}) and the type of energy tariff for electricity. So, we have:

$$y'^*_i = x_i'\beta_1 + P_{2012_i}\beta_2 + \varepsilon_i$$

with

$$P_{2012_i} = \gamma_1 P_{2011_i} + \gamma_2 TARIFF + v_{i,k}$$

Where y'* the final energy consumed and x_i the regressors.

We estimate the model using a double least squares model, which enables us to correct for the endogeneity issue of enrgy prices. In addition, two-step methods can lead to an underestimation of standard errors of the second step. We apply a bootstrap correction on the variance-covariance matrix to avoid bias in the interpretation of coefficients' significance level.

4.2 Quantile regressions

The second analysis focuses on determinants of extreme behaviors regarding energy use. By using quantile regressions, we could precise the differentiated impacts of socio-economic determinants on several energy consumption levels (under consumption, normal consumption and over consumption), which are estimated by the *Intensity of Energy Use* indicator. The quantile regression method is an extension of ordinary regression[10]. It was first introduced by Koenker and Basset (Koenker and Bassett 1978) and generalizes the concept of univariate quantile to a conditional quantile given one or more covariates. So, it is less restrictive than the OLS method because slope coefficients can vary across the chosen quantiles of the dependent variable and so, are not only mean estimations. This method allows to detect if explanatory determinants have the same effects or not for extreme values of the dependent variable (for example for 5^{th}, 25^{th} and 75^{th} quantiles) and to quantify these effects. In addition to giving robust coefficients estimations with respect to outliers, in our case, it is also useful to assess the variability of the main determinants of over and under-consumption, behaviors represented by extreme values of our dependent variable *Intensity of energy use.* By doing so, we may detect differential impacts of revenue, energy price or individual data such as preferences depending of the level of consumption. As an example, the research of N. Kaza (International Energy Agency (IEA)) used this method to estimate the impacts of numerous determinants on different quantiles of energy consumption in the residential sector in the US (Kaza 2010). It showed that neighborhood density, housing size and housing type effects on the tails of the distribution are substantially different.

[10] We ensure before choosing the quantile regression the absence of the energy price endogeneity with the dependent variable. Using the same tests than those presented in appendix C3, we confirm the validity of instruments and the absence of endogeneity.

So we have the θ^{th} quantile is generally defined as:

$q_\theta(Y) = inf\{y: F_Y(y) \geq \theta\}, 0 < \theta < 1$

The linear quantile regression in our model can be formalized as follows:
$y_i = x_i'\beta_\theta + u_i$ with $Q_\theta(y_i \ / \ x_i) = x_i' \ \beta_\theta$

In our model y is the vector of *Intensity of energy use* data (in logarithm), x is a vector of all the regressors, β is the vector parameters to be estimated, and u is a vector of residuals. $Q_\theta(y_i \ / \ x_i)$ is the Q^{th} quantile of y_i given x_i. The θ^{th} QR estimator minimizes over $\widehat{\beta_\theta}$ the objective function (Cameron and Trivadi 2010):

$$\mathrm{Q}(\beta_\theta) = \sum_{i:yi \geq x_i'\beta}^{N} \theta|y_i - x_i'\beta_\theta| + \sum_{i:yi < x_i'\beta}^{N} (1-\theta)|y_i - x_i'\beta_\theta|$$

If $\theta = 0.75$, much more weight is placed on prediction for observations with $y \geq x'\beta$ than for observation with $y < x'\beta$. Finally, asymptotic and bootstrapping methods are normally used to obtain the standard errors and confidence limits for coefficient estimates.

For or model we first use quantiles 0.05 0.25 0.57 0.75 0.9 to distinguish the specific effects of determinants on extreme behaviors (under consumption IEU<0.33 and over consumption IEU> 1.75) and then we represent graphically the evolution of coefficients at each quantile. Quantile 0.57 represents an IEU equal to 1. Finally, quantiles 0.25 and 0.75 respectively represents an IEU equals to 0.65 and 1.31.

5. Results

5.1 Drivers of energy consumption: Discrete-continuous choice model

Results of the two-steps are presented in Table 5 below.

Table 4: Results of discrete-continuous model

VARIABLES	(International Energy Agency (IEA)) Discrete choice - Results for ordered logit	(2) Continuous Choice - IV 2SLS regression 1	(3) Continuous Choice - IV 2SLS regression 2
Energy price in 2012 (log)		-1.955***	-2.002***
		(0.518)	(0.500)
Income (log)	-0.260***	0.0980***	0.100***
	(0.0977)	(0.0344)	(0.0348)
Number of persons		0.0675***	0.0659***
		(0.0138)	(0.0133)
Inferior to 44 years old	Ref	Ref	Ref
Age between 44 and 55 years	-0.0383	0.0222	0.0236
	(0.124)	(0.0346)	(0.0351)
Age between 56 and 66 years	0.273**	-0.0166	-0.0101
	(0.129)	(0.0440)	(0.0452)
More than 66 years old	0.590***	-0.0905	-0.0861
	(0.143)	(0.0572)	(0.0589)
tenant	0.339***	0.00102	-0.00237
	(0.120)	(0.0335)	(0.0337)
preference for comfort for heating	0.0346	0.0749***	
	(0.0843)	(0.0247)	
Number of days of housing vacancy during heating period (log)		-0.0280***	-0.0255***
		(0.00983)	(0.00950)
Number of appliances (log)		0.156***	0.143***
		(0.0376)	(0.0353)
Heating temperature			0.0443***
			(0.00938)
Surface of the dwelling (log)	-1.959***	0.175***	0.179***
	(0.145)	(0.0634)	(0.0642)
Climate zone H3	Ref	Ref	Ref
Climate zone H2	-1.053***	0.145***	0.163***
	(0.0929)	(0.0428)	(0.0431)
Climate zone H1 - coldest	-1.185***	0.0684	0.0594
	(0.187)	(0.0716)	(0.0715)
Town < 2000 inhabitants	0.607***	-0.180***	-0.189***
	(0.171)	(0.0563)	(0.0556)
Town between 2,000 and 10,000 inhabitants	0.354*	-0.0628	-0.0810
	(0.186)	(0.0545)	(0.0528)
Town between 10,000 and 50,000 inhabitants	0.0382	-0.0852	-0.0952*
	(0.181)	(0.0533)	(0.0527)
Town between 50,000 and 200,000 inhabitants	-0.266	0.0226	0.0209
	(0.195)	(0.0545)	(0.0510)
City between 200,000 and 2,000,000 inhabitants	-0.152	-0.0326	-0.0472

VARIABLES	(International Energy Agency (IEA)) Discrete choice - Results for ordered logit	(2) Continuous Choice - IV 2SLS regression 1	(3) Continuous Choice - IV 2SLS regression 2
	(0.169)	(0.0470)	(0.0443)
Paris	Ref	Ref	Ref
Predicted value for class A-B		-1.886	-1.790
		(1.584)	(1.646)
Predicted value for class C		-2.947***	-2.881***
		(0.708)	(0.745)
Predicted value for class D		-2.011**	-1.979**
		(0.811)	(0.808)
Predicted value for class E		-1.283***	-1.243***
		(0.463)	(0.471)
Predicted value for class F		-1.686	-1.600
		(1.147)	(1.141)
Predicted value for class G			
		Ref	Ref
Duration since move in	0.0208***		
	(0.00347)		
Non-detached house	-0.218**		
	(0.0995)		
Individual housing unit	0.782***		
	(0.148)		
Constant cut1	-15.15***		
	(1.052)		
Constant cut2	-12.90***		
	(1.036)		
Constant cut3	-11.20***		
	(1.027)		
Constant cut4	-9.639***		
	(1.020)		
Constant cut5	-8.474***		
	(1.018)		
Constant		4.387***	3.502***
		(0.474)	(0.510)
Observations	2,044	2,044	2,044
R-squared		0.159	0.166

*Note: Robust standard errors in parentheses *** p<0.01, ** p<0.05, ***p<0.1*

5.1.1 Ordered logit (A-B is the reference class)

Income has a significant negative effect: households with higher revenue are more likely to live in energy efficient homes than less richer households. Low income households have a relatively smaller opportunity to live in an energy efficient dwelling. This result is in line with

Santamouris et al. (2007). A positive elasticity may involve mainly the purchase or the rent of more efficient dwelling. The age of the reference person is also impacting: for the two higher age classes, households are more likely to live in a non-efficient dwelling than before 44, the effect being higher for households in the last category (more than 66 years). Concerning occupancy characteristics, we can note that tenants are more likely to live in non-energy efficient homes than owners (Similar results are found in Charlier, 2015 #1921). It could be explained by the split incentives issue: rented dwellings are more likely to be theoretically non-efficient because owners would not benefit from energy savings won thanks to thermal retrofit works. Moreover, dwelling occupancy period has a significant link with energy performance of the dwelling. The more recent the move in date is, the more likely to live in efficient dwelling households are. Two assumptions can be made: the higher availability of energy efficient dwellings on the current real estate market and/or the higher attention paid by households to dwelling energy information (since several years, EPC information is provided when renting or buying a dwelling). Some environmental characteristics are also correlated to energy performance level of dwellings. Concerning neighborhood, the less isolated the dwelling (in terms of shared walls) is, the more energy efficient it is likely to be. Urban area types are also impacting; compared to Paris and big cities, dwellings in rural area are more likely to be non-energy efficient (Belaïd 2016). Moreover, energy efficient dwellings are more likely to be found in cooler climate zones. Finally, size and building type effects are also significant; the bigger the dwelling is, the more energy efficient it is likely to be; living in a bigger house increases the probability of being in an energy efficient dwelling.

Finally, households' preferences have no impact on the decision to live in an energy efficient dwelling.

5.1.2 Second step: individual households' characteristics role in energy consumption variability

Introducing predicted probabilities given by the first step of our two-stage into the continuous choice model allows us to capture interactions between energy efficiency levels of dwellings and its environmental and occupancy characteristics while controlling energy performance of the dwelling. Thus, we extract from the second stage direct effects of individual characteristics on energy consumption variability. To control the quality of instruments, different tests are performed. Results are presented in appendix C1. We can note that income effect remains significantly positive in comparison to results provided by our literature review. It means that, the higher its income is, the more the household is going to consume energy (all uses combined).

The variables representing the age of the reference person are not significant, meaning that major effects of household age is that which results from dwelling choice. As well, tenure status has no significant effect. Moreover, the effect on energy consumption of appliances number is significant and positive, as expected. In terms of energy price elasticity, the effect is also significant, negative and important. Price elasticity is almost -2.0.

Concerning behavioral and preferences determinants, our model shows an original result: individual preferences for comfort over economy for heating use are highly significant and have a direct positive effect on energy consumption, meaning that a part of energy consumption could not be targeted by public policies aiming to reduce energy. This positive effect is confirmed thanks to regression 3, showing, a significant positive effect of heating temperatures on energy consumption. Regarding behavioral data, we see that, the duration of day-vacancy has unsurprisingly a negative significant effect on total energy consumption.

Finally, surface has a direct positive effect on energy consumption: the bigger the household is, the more energy by square meters it will consume. Predicted probabilities controlling for theoretical energy performance and interactions has consistent effects, the more efficient the home is, the less we consume. However, energy class E is less efficient as expected (compared to F and D), it could be due to rebound effect.

5.2 Understanding the energy performance gap: the quantile regression

By using a quantile regression, we focus on determinants of extreme behaviors regarding energy consumption. Results show that impacts of selected determinants vary across IEU levels. Results are presented in appendix C2 in Table (10-11-12). Moreover, proof of the absence of endogeneity due to energy prices are also provided in appendix C3.

Income affects more “over consumption” than “normal behavior” and “under consumption”. Two interpretations can be given: higher income are more likely in energy efficient class (see step 1 from first model) and/or high energy class implies bigger energy performance gap in the sense of overconsumption. Here, income elasticities are positive and are in line with a great part of economic literature on energy demand (Labandeira et al. 2006); Nesbakken (1999), 2001; Risch and Salmon 2017).

Regarding effect of energy price, we can note that price elasticity is more elastic on lower intensity of energy use (under consumption). Households in these dwellings are more likely to be in non-energy efficient dwellings (see coefficients of energy class in quantile regression), to be poorer (step 1 of first model) and may be then more sensitive to price variability. The price elasticities are between -0.375 and -0.986, which is slightly higher than the results obtained for the demand price elasticity (Dubin and McFadden 1984; R. Halvorsen 1975; Larsen and Nesbakken 2004; Nesbakken 1999, 2001; Risch and Salmon 2017).

Renter are supposed to have bigger intensity of energy use than owner. This effect is positive for low values of intensity of energy use. But with our first model, we saw that they are more likely to live in non-energy efficient dwelling.

Age class of the reference person has a positive effect (compared to the class under 44 years) but is mostly non-significant except for q0.75 and q0.25.

Regarding preference for comfort over economy for heating, there is a positive effect always significant (for the 3 given quantiles). Results show that preferences have a growing impact with higher value of the intensity of energy use indicator: effect of individual preferences is then growing with higher IEU. If we consider preferences in level (low, medium and high level, see Appendix C2), we note that strongly preferences for comfort are associated with a higher IEU. 12% of the total variance in IEU is explained by preferences variables, which represent an unneglectable share.

Number of appliances has a positive impact and bigger impact for low intensity of energy use values. People who force themselves to consume less (q0.25) may reduce their use of appliances.

Heating temperature: growing effect with higher value of intensity. Significant for high value of the indicator. Increase of temperature is significantly correlated with overconsumption

Concerning impact of theoretical energy performance on IEU levels, we see that the higher the energy class is, the higher the effect on intensity of energy use is. Moreover, the effect is more important with higher value of the indicator: +0.8 on high values of the indicator when living in A or B energy class. This could constitute a proof for comfort effect, usually called “rebound effect”.

Environmental characteristics also have a differentiated impact on IEU levels that can be mentioned. The fact to belong to climate zone 2 is positively correlated with energy over-consumption behaviors. We saw earlier that the probability to live in an energy efficient dwelling was higher when living in colder climate zone than the zone 1 and that a household living in an energy efficient dwelling was more likely to have an over-consumption behavior. So, it could mean that high IEU and cold climate zone are indirectly linked via higher comfort effect. Finally, the size of the urban area has an ambivalent impact on IEU levels; in comparison to Paris, living in small towns and cities has a positive and growing significant effect on high levels of IEU; but from 50 000 to 2 000 000 inhabitants, the effect becomes slightly negative (still compared to Paris). We can assume that households in small towns or cities do not benefit from neighborhood heating and must heat more their homes for the same level of comfort. This "small cities" effect is more impacting on high IEU levels.

Finally, the model allows us to valid the stated assumption: the energy performance gap is partly driven by individual characteristics (socio-economic, preferences for comfort and behavior). .

6. Conclusion and policy implications

This research provides a new proof of the significant role of individual characteristics on energy consumption variability. Common contributors to total energy consumption and high intensity of energy use confirmed in this paper are the followings: income, number of persons in the dwellings, age of the reference person, number of appliances and preferences for comfort over economy concerning heating. Moreover, we have seen that the preferences for comfort hold a place in the understanding of energy consumption variability, especially because of its significant link with IEU, indicating that over-consumption behaviors are partly driven by households' heterogeneity. Additionally, we note that energy efficiency of dwellings are factors also associated with over-consumption behaviors.

In terms of public policies, this research has two contributions.

First, by highlighting the role of individual preferences on energy consumption, its link with energy over-consumption and energy efficiency of the dwelling, we warn public authorities of the risk of non-reaching energy savings expected after thermal retrofits because of households' heterogeneity and comfort effect. If public funds were committed to finance energy retrofits, there is so a risk of the loss of efficiency concerning their allocation. Indeed, we saw that preferences for comfort for heating had an important impact on high levels of IEU, which are also associated with higher theoretical energy performances of dwellings. As a result, retrofits renovations (aiming at improving energy efficiency) could imply changes in preferences (in the sense of an increasing of preferences for comfort) and then an increasing of over-consumption.

Second, as the part of energy consumption variability due to individual characteristics is clearly significant, we recommend users of prospective tools on energy consumption to find a way of integrating heating behaviors for more high realism. More precisely, they should include the fact that people living in non-energy efficiency dwellings restrict their energy consumption relatively to what it is theoretically expected (appendix B), in opposition to people living in energy-efficient dwellings, who tend to over-consume. So, it is necessary to carefully consider post retrofits heating behaviors and thermal indoor temperature levels when assessing energy benefits of energy retrofits, especially when using conventional measures of energy performance to predict energy consumption. Results may be less ambitious than expected. Finally, based on the results of this research, policies could think on new tools to reduce the part of the gap caused by consumer heterogeneity such as nudges, which are based on consumer environmental sensibility.

References

A. Greening, Lorna, Greene, David L., and Difiglio, Carmen (2000), 'Energy efficiency and consumption -- the rebound effect -- a survey', *Energy Policy,* 28 (6-7), 389-401.

Allibe, Benoît (2012), 'Modélisation des consommations d'énergie du secteur résidentiel français à long terme - Amélioration du réalisme comportemental et scénarios volontaristes', (Ecole des Hautes Etudes en Sciences Sociales (EHESS)).

Aydinalp, Merih, Ismet Ugursal, V., and Fung, Alan S. (2002), 'Modeling of the appliance, lighting, and space-cooling energy consumptions in the residential sector using neural networks', *Applied Energy,* 71 (2), 87-110.

Basmann, R.L (1960), 'On finite sample distributions of generalized classical linear identifiability test statistics.', *Journal of the American Statistical Association,* 55, 650–59.

Belaïd, Fateh (2016), 'Understanding the spectrum of domestic energy consumption: Empirical evidence from France', *Energy Policy,* 92, 220-33.

Belaïd, Fateh and Garcia, Thomas (2016), 'Understanding the spectrum of residential energy-saving behaviours: French evidence using disaggregated data', *Energy Economics,* 57, 204-14.

Berkhout, Peter H. G., Muskens, Jos C., and W. Velthuijsen, Jan (2000), 'Defining the rebound effect', *Energy Policy,* 28 (6-7), 425-32.

Bourrelle, Julien S. (2014), 'Zero energy buildings and the rebound effect: A solution to the paradox of energy efficiency?', *Energy and Buildings,* 84 (0), 633-40.

Brounen, Dirk and Kok, Nils (2011), 'On the economics of energy labels in the housing market', *Journal of Environmental Economics and Management,* 62 (2), 166-79.

Brounen, Dirk, Kok, Nils, and Quigley, John M. (2013), 'Energy literacy, awareness, and conservation behavior of residential households', *Energy Economics,* 38 (0), 42-50.

Cameron, Colin A. and Trivadi, Pravin K. (2010), *Microeconometrics Using Stata* (Stata Press, Revisited Version).

Carassus, Jean, et al. (2013), 'Performances environnementales de l'immobilier : du conventionnel au réel'.

Cayla, Jean-Michel, Maizi, Nadia, and Marchand, Christophe (2011), 'The role of income in energy consumption behaviour: Evidence from French households data', *Energy Policy,* 39 (12), 7874-83.

Charlier, Dorothée (2015), 'Energy efficiency investments in the context of split incentives among French households', *Energy Policy,* 87, 465-79.

Dubin, Jeffrey A. and McFadden, Daniel L. (1984), 'An econometric analysis of residential electric appliance holdings and consumption', *Econometrica,* 52 (2), 345-62.

Estiri, Hossein (2014), 'Building and Household X-Factors and Energy Consumption at the Residential Sector: A Structural Equation Analysis of the Effects of Household and Building Characteristics on the Annual Energy Consumption of US Residential Buildings', *Energy Economics,* 43, 178-84.

--- (2015), 'The indirect role of households in shaping US residential energy demand patterns', *Energy Policy,* 86, 585-94.

Ewing, Reid and Rong, Fang (2008), 'The impact of urban form on U.S. residential energy use', *Housing Policy Debate,* 19 (1), 1-30.

Frondel, Samuel, Flores, Fernanda Martínez , and Vance, Colin (2016), 'Heterogeneous Rebound Effects: Comparing Estimates from Discrete-Continuous Models', *USAEE Working Paper 16* (Ruhr Economic Papers: RWI - Leibniz-Institut für Wirtschaftsforschung, Ruhr-University Bochum, TU Dortmund University, University of Duisburg-Essen).

Galvin, Ray (2010), 'Thermal upgrades of existing homes in Germany: The building code, subsidies, and economic efficiency', *Energy and Buildings,* 42 (6), 834-44.

--- (2014a), 'Making the 'rebound effect' more useful for performance evaluation of thermal retrofits of existing homes: Defining the 'energy savings deficit' and the 'energy performance gap'', *Energy and Buildings,* 69, 515-24.
--- (2014b), 'Estimating broad-brush rebound effects for household energy consumption in the EU 28 countries and Norway: some policy implications of Odyssee data', *Energy Policy,* 73, 323-32.
--- (2015), 'How many interviews are enough? Do qualitative interviews in building energy consumption research produce reliable knowledge?', *Journal of Building Engineering,* 1 (0), 2-12.
Galvin, Ray and Sunikka-Blank, Minna (2013), 'Economic viability in thermal retrofit policies: Learning from ten years of experience in Germany', *Energy Policy,* 54, 343-51.
--- (2014), 'Disaggregating the causes of falling consumption of domestic heating energy in Germany', *Energy Efficiency,* 7 (5), 851-64.
Halvorsen, Bente and Larsen, Bodil M. (2001), 'The flexibility of household electricity demand over time', *Resource and Energy Economics,* 23 (1), 1-18.
Halvorsen, Robert (1975), 'RESIDENTIAL DEMAND FOR ELECTRIC ENERGY', *Review of Economics & Statistics,* 57 (1), 12-18.
Hausman, J. A (1978), ' Specification tests in econometrics', *Econometrica,* 46 (1251–1271).
Heckman, James J. and Robb, Richard (1985), 'Alternative methods for evaluating the impact of interventions', *Journal of Econometrics,* 30 (1), 239-67.
International Energy Agency (IEA), OECD, OPEC, WB, 2010 (2010), 'Analysis of the Scope of Energy Subsidies and Suggestions for the G20 Initiative', (Joint Report, Toronto).
Jiang, Lei, Folmer, Henk, and Ji, Minhe (2014), 'The drivers of energy intensity in China: A spatial panel data approach', *China Economic Review,* 31 (0), 351-60.
Jones, Rory V., Fuertes, Alba, and Lomas, Kevin J. (2015), 'The socio-economic, dwelling and appliance related factors affecting electricity consumption in domestic buildings', *Renewable and Sustainable Energy Reviews,* 43, 901-17.
Jones, Rory V., et al. (2016), 'Space heating preferences in UK social housing: A socio-technical household survey combined with building audits', *Energy and Buildings,* 127, 382-98.
Kaza, Nikhil (2010), 'Understanding the spectrum of residential energy consumption: A quantile regression approach', *Energy Policy,* 38 (11), 6574-85.
Koenker, Roger and Bassett, Gilbert Jr. (1978), 'Regression Quantiles', *Econometrica,* 46 (1), 33-50.
Kriström, Bengt (2006), 'Household Behaviour and the Environment Reviewing the Evidence ', (OECD: https://www.oecd.org/environment/consumption-innovation/42183878.pdf).
Labandeira, Xavier, Labeaga, JosÃ© M., and Rodriguez, Miguel (2006), 'A residential energy demand system for Spain (English)', *The Energy journal (Cambridge, MA),* 27 (2), 87-111.
Larsen, Bodil Merethe and Nesbakken, Runa (2004), 'Household electricity end-use consumption: results from econometric and engineering models', *Energy Economics,* 26 (2), 179-200.
Leahy, Eimear and Lyons, Sean (2010), 'Energy use and appliance ownership in Ireland', *Energy Policy,* 38 (8), 4265-79.
Lopes, M. A. R., Antunes, C. H., and Martins, N. (2012), 'Energy behaviours as promoters of energy efficiency: A 21st century review', *Renewable and Sustainable Energy Reviews,* 16 (6), 4095-104.
Majcen, D., Itard, L. C. M., and Visscher, H. (2013), 'Theoretical vs. actual energy consumption of labelled dwellings in the Netherlands: Discrepancies and policy implications', *Energy Policy,* 54, 125-36.
McLoughlin, Fintan, Duffy, Aidan, and Conlon, Michael (2012), 'Characterising domestic electricity consumption patterns by dwelling and occupant socio-economic variables: An Irish case study', *Energy and Buildings,* 48, 240-48.
Nesbakken, Runa (1999), 'Price sensitivity of residential energy consumption in Norway', *Energy Economics,* 21 (6), 493-515.
--- (2001), 'Energy Consumption for Space Heating: A Discrete-Continuous Approach', *Scandinavian Journal of Economics,* 103 (1), 165–84.

Rehdanz, Katrin (2007), 'Determinants of residential space heating expenditures in Germany', *Energy Economics,* 29 (2), 167-82.
Risch, Anna and Salmon, Claire (2017), 'What matters in residential energy consumption: evidence from France', *International Journal of Global Energy Issues,* 40 (1-2), 79-115.
Santamouris, M., et al. (2007), 'On the relation between the energy and social characteristics of the residential sector', *Energy and Buildings,* 39 (8), 893-905.
Sargan, J.D (1958), 'The estimation of economic relationships using instrumental variables', *Econometrica,* 26, 393–415.
Song, Nianfu, et al. (2012), 'Factors affecting wood energy consumption by U.S. households', *Energy Economics,* 34 (2), 389-97.
Sorrell, Steve and Dimitropoulos, John (2008), 'The rebound effect: Microeconomic definitions, limitations and extensions', *Ecological Economics,* 65 (3), 636-49.
Stock, J. H and Yogo, M. (2005), 'Testing for weak instruments in linear IV regression', in D. W. K. Andrews and J. H. Stock (ed.), *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenber* (New York: Cambridge University Press), 80-108.
Thomas, Brinda A. and Azevedo, Inês L. (2013a), 'Estimating direct and indirect rebound effects for U.S. households with input–output analysis. Part 2: Simulation', *Ecological Economics,* 86 (0), 188-98.
--- (2013b), 'Estimating direct and indirect rebound effects for U.S. households with input–output analysis Part 1: Theoretical framework', *Ecological Economics,* 86 (0), 199-210.
Vaage, Kjell (2000), 'Heating technology and energy use: a discrete/continuous choice approach to Norwegian household', *Energy Economics,* 22 (6), 649.
Wirl, Franz (1987), 'Thermal Comfort, Energy Conservation and Fuel Substitution: An Economic-Engineering Approach ', *Energy System Policy,* 114 (January).
Wu, D.M (1974), 'lternative tests of independence between stochastic regressors and disturbances: Finite sample results', *Econometrica,* 42, 529–46.
Yohanis, Yigzaw Goshu (2012), 'Domestic energy use and householders' energy behaviour', *Energy Policy,* 41 (0), 654-65.

Appendix

A. Energy prices

Table 6: Energy prices provided by PEGASE database

	2011	2012
ELECTRICITY TARIFF		
Electricity, blue tariff, base option in euros (tax included)		
Annual subscription cost 3 kVA	64,94606	67,40325
Annual subscription cost 6 kVA	77,45169	80,36592
Annual subscription cost 9 kVA	90,3377	93,76717
Annual subscription cost 12 kVA	142,84527	148,13392
Annual subscription cost 15 kVA	164,85725	171,04758
Annual subscription cost 18 kVA	219,2238	227,44092
Price for 100 kWh (power 3 kVA)	17,02237	17,7994
Price for 100 kWh (power 6 kVA)	16,23193	16,9816
Electricity, blue tariff, peak hours tariff in euros (tax included)		
Annual subscription cost 6 kVA	93,13223	96,59658
Annual subscription cost 9 kVA	111,76704	115,91475
Annual subscription cost 12 kVA	189,49559	196,56458
Annual subscription cost 15 kVA	223,04773	231,32342
Annual subscription cost 18 kVA	254,38013	263,81675
Annual subscription cost 24 kVA	529,87303	549,78758
Annual subscription cost 30 kVA	652,50116	677,02358
Annual subscription cost 36 kVA	754,42164	782,73067
100 kWh peak-hours	12,91385	13,54292
100 kWh peak-off	8,76965	9,23933
Price for 100 kWh (power 6 kVA)	14,03546	14,70435
Price for 100 kWh (power 9 kVA)	13,02266	13,65389
Price for 100 kWh (power 12 kVA)	12,77758	13,39973
Electricity, blue tariff, tempo option in euros (tax included)		
Annual subscription cost 9 kVA	109,04157	113,022
Annual subscription cost 12 kVA	203,35865	210,90942
Annual subscription cost 30 kVA	456,64613	473,54025
Annual subscription cost 36 kVA	566,42158	587,43975
100 kWh blue days and peak-off	6,8142	7,2111
100 kWh blue days and peak-hours	8,20155	8,65528
100 kWh white days and peak-off	9,8401	10,35061
100 kWh white days and peak-hour	11,7537	12,33594
100 kWh red days and peak-off	18,5589	19,40033
100 kWh red days and peak-hour	49,16455	51,17409
Electricity, market tariff, in euros (tax included)		
All tariff	13,41974	13,82434
Tariff DA	24,45679	25,13133
Tariff DB	15,8404	16,3847
Tariff DC	14,02566	14,45913

Tariff DD	12,84391	13,2134
Tariff DE	12,54369	12,91665
GAS TARIFF		
Natural Gas, price in euros (tax included)		
Annual subscription cost - base tariff	43,8933	46,92645
Annual subscription cost - tariff B0	58,0092	61,97075
Annual subscription cost - tariff B1	185,18415	195,4546
Annual subscription cost -tariff B2I	185,18415	195,4546
100 kWh PCS - base tariff	9,3988	9,96987
100 kWh - tariff B0	8,0742	8,51871
100 kWh- tariff B1	5,58353	5,86163
100 kWh - tariff B2I	5,58353	5,86163
Price for 100 kWh tariff B0	11,74238	12,42551
Price for 100 kWh tariff B1	7,08853	7,44654
Price for 100 kWh tariff B2I	6,79365	7,13536
DOMESTIC OIL TARIFF		
Tariff of one ton of propane in tank	1670,297	1791,087
100 kWh PCI (Lower calorific value) propane in tank	12,96815	13,90596
Price of one ton of propane	1670,297	1791,087
100 kWh PCS (Higher calorific value) of propane	12,1036	12,97889
100 kWh PCI of propane	13,06961	14,01476
Bottle of 13 kg butane	30,19	31,75
100 liters of oil at Rate C1	88,79	96,88
100 kWh oil PCI at Rate C1	8,90482	9,71618
WOOD TARIFF		
One ton of bulk pellets	250	260
A Sire of Logs	63	67
100 kWh PCI of bulks	3,70588	3,94118

Source: PEGASE database, French Ministry of Energy

B. Descriptive statistics – Energy performance class

B.1 Main descriptive statistics

Table 7 : Main descriptive statistics

Variables	Obs	Mean	Std. Dev.	Min	Max
Income	2044	40396.59	24628.58	307	277601
Number of persons	2044	2.543053	1.297086	1	10
Age of the reference person : between 45 and 55 years	2044	0.2372798	0.4255193	0	1
Age of the reference person : between 56 and 65 years	2044	0.2632094	0.4404828	0	1
Age of the reference person : over 66 years	2044	0.2514677	0.4339629	0	1
Tenant	2044	0.2245597	0.4173941	0	1
Preference for comfort over economy : heating	2044	0.5650685	0.4958693	0	1
Preference for comfort over economy : electricity	2044	0.4124266	0.4923916	0	1
Preference for comfort over economy : hot water	2044	0.5552838	0.4970559	0	1
Heating temperature	2044	19.9147	1.460411	8	30
Number of days of housing vacancy during heating period	2044	7.781311	16.08995	0	210
Number of appliances	2044	16.08023	13.50374	1	341
Climate zone 2	2044	0.3610568	0.4804245	0	1
Climate zone 3	2044	0.0631115	0.243223	0	1
Town < 2000 inhabitants	2044	0.2739726	0.4461042	0	1
Town between 2,000 and 10,000 inhabitants	2044	0.1320939	0.3386757	0	1
Town between 10,000 and 50,000 inhabitants	2044	0.1482387	0.3554234	0	1
Town between 50,000 and 200,000 inhabitants	2044	0.109589	0.3124533	0	1
City between 200,000 and 2,000,000 inhabitants	2044	0.2426614	0.4287969	0	1
IEU	2044	1.050385	0.6290751	.010	5.29
IEU (gas sample)	573	0.8289124	0.4720684	0	4.13
Energy price 2012	2044	0.0867332	0.0360352	0	.381
Energy price 2011	2044	0.0821862	0.0343162	0	.31
Time since last move-in date (years)	2044	17.26174	15.01292	0	89
Surface	2044	111.6899	49.03991	15	600
Final energy consumption (kwh/m2/year)	2044	171.9225	106.0669	2.26	1401.52
Theoretical energy consumption(kwh/m2/year)	2044	203.5162	123.2094	12.73	1444.58

B.2 Energy class

Figure 4: EPC energy classes

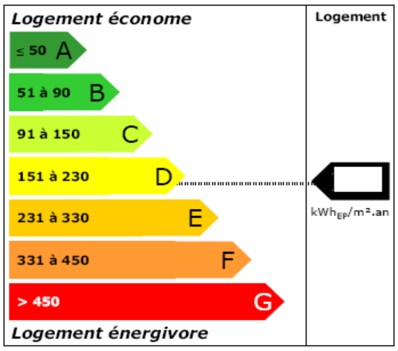


Table 8: Distribution of national dwelling stock into energy classes

Energy Class	Number of observations	At national scale	Share of the housing stock (%)
A-B	48	439 585	2
C	281	2 724 895	12.6
D	564	5 483 573	25.4
E	602	6 322 821	28.3
F	301	3 361 569	15.6
G	250	3 257 038	15

B.3 Assessment of the energy performance gap

Table 5: Heating temperature and IEU

IEU (4 quantiles, weighted)	IEU<0.65	0.65<IEU<0.92	0.92<IEU<1.31	1.31<IEU
Heating temperature	19.6 (1.6)	19.8 (1.3)	20.1 (1.4)	20.2 (1.5)
Number of observations	487	517	524	517

C. Regressions

C.1 Quality test of instruments

First, we can perform test to determine whether endogenous regressors in the model are in fact exogenous. After 2SLS estimation with an unadjusted VCE, the Durbin (Jiang et al. 2014) and Wu–Hausman (Hausman 1978; Wu 1974) statistics are reported.

We control consistency of the results with a VCE estimation. In all cases, if the test statistic is significant, then the variables being tested must be treated as endogenous.

Table 6: test of endogeneity

Ho: variables are exogenous	
Durbin (score) chi2(International Energy Agency (IEA)) = 3.87242	p = 0.0491
Wu-Hausman F(1,2019)= 3.83232	p = 0.0504
Robust score chi2(International Energy Agency (IEA)) = 3.49652	p = 0.0615
Robust regression F(1,2019) = 2.43238	p = 0.1190

Second, we now explore the degree of correlation between the additional instruments (energy prices in 2011 and electricity energy tariffs) and the endogenous regressor (energy prices in 2012). Our exogenous variable can be considered as a valid instrument, if it is correlated with the included endogenous regressors but uncorrelated with the error term. Using a Stock and Yogo (2005) test, we can discuss the validity of the instruments. The null hypothesis of each of Stock and Yogo's tests is that the set of instruments is weak. To perform these Wald tests, we choose a relative rejection rate of 5%. If the test statistic exceeds the critical value, we can conclude that our instruments are not weak. In our model, the F statistic is 89822.3 and largely exceeds the critical value. Our instruments are not weak.

Minimum eigenvalue statistic = 89822.3

	5%	10%	20%	30%
2SLS relative bias	13.91	9.08	6.46	5.39
2SLS Size of nominal 5% Wald test	22.30	12.83	9.54	7.80
LIML Size of nominal 5% Wald test	6.46	4.36	3.69	3.32

Finally, to confirm our results, we perform tests of overidentifying restrictions. With the 2SLS estimator, Sargan's (Sargan 1958) and Basmann's (Basmann 1960) χ^2 tests are reported. A statistically significant test statistic always indicates that the instruments may not be valid. Here, tests are not significant. Instruments are valid.

Sargan (score) chi2(2) = 0.544109	p = 0.7618
Basmann chi2(2) = 0.537331	p = 0.7644

C.2 Quantile regressions

Table 10: Results of Quantile regression with preference for comfort for heating

	(International Energy Agency (IEA)) OLS regression	(2) Quantile regressions				
VARIABLES		0.05	0.25	0.57	0.75	0.90
Energy price in 2012 (log)	-0.729***	-0.765***	-0.578**	-0.336	-0.661**	-0.596**
	(0.173)	(0.290)	(0.233)	(0.205)	(0.262)	(0.274)
Income (log)	0.0323**	0.0201	0.0168	0.0377**	0.0479**	0.0120
	(0.0134)	(0.0237)	(0.0149)	(0.0160)	(0.0218)	(0.0227)
Number of persons	0.0192***	0.0190**	0.0168**	0.0188**	0.0167**	0.0155
	(0.00518)	(0.00887)	(0.00661)	(0.00740)	(0.00835)	(0.0105)
Age inferior to 44 years old	Ref	Ref	Ref	Ref	Ref	Ref
Age between 44 and 55 years	-0.000200	-0.00136	0.00611	0.00171	0.000185	-0.0300
	(0.0144)	(0.0262)	(0.0186)	(0.0188)	(0.0222)	(0.0300)
Age between 56 and 66 years	0.0355**	-0.00592	0.0236	0.0254	0.0445*	0.0633**
	(0.0165)	(0.0314)	(0.0199)	(0.0207)	(0.0252)	(0.0314)
More than 66 years old	0.0298*	0.00345	0.0443**	0.0279	0.0218	0.00761
	(0.0170)	(0.0318)	(0.0216)	(0.0225)	(0.0257)	(0.0352)
tenant	-0.0111	0.000275	0.00426	-0.00998	-0.00952	-0.0430
	(0.0134)	(0.0259)	(0.0168)	(0.0162)	(0.0217)	(0.0282)
number of days of housing vacancy during heating period (log)	-0.00691*	-0.00350	-0.00597	-0.00599	-0.00674	-0.00896
	(0.00375)	(0.00823)	(0.00496)	(0.00517)	(0.00541)	(0.00727)
Number of appliances (log)	0.0569***	0.0656***	0.0765***	0.0598***	0.0546**	0.0487
	(0.0152)	(0.0212)	(0.0155)	(0.0200)	(0.0216)	(0.0356)
preference for comfort for heating	0.0310***	-0.00462	0.0394***	0.0288**	0.0428***	0.0638***
	(0.0102)	(0.0188)	(0.0121)	(0.0122)	(0.0153)	(0.0202)
Class AB	0.667***	0.480***	0.502***	0.611***	0.830***	1.072***
	(0.0586)	(0.0782)	(0.0829)	(0.0582)	(0.145)	(0.131)
Class C	0.426***	0.330***	0.414***	0.426***	0.446***	0.482***
	(0.0227)	(0.0445)	(0.0282)	(0.0258)	(0.0309)	(0.0492)
Class D	0.290***	0.243***	0.275***	0.283***	0.321***	0.388***
	(0.0185)	(0.0341)	(0.0251)	(0.0197)	(0.0274)	(0.0340)
Class E	0.182***	0.148***	0.183***	0.182***	0.212***	0.215***
	(0.0167)	(0.0361)	(0.0237)	(0.0176)	(0.0242)	(0.0300)
Class F	0.0872***	0.0558*	0.107***	0.101***	0.0859***	0.109***
	(0.0180)	(0.0329)	(0.0279)	(0.0181)	(0.0247)	(0.0324)
Class G	Ref	Ref	Ref	Ref	Ref	Ref
Surface of the dwelling (log)	-0.174***	-0.107***	-0.144***	-0.181***	-0.186***	-0.211***
	(0.0164)	(0.0350)	(0.0198)	(0.0192)	(0.0247)	(0.0294)
Climate zone H3	Ref	Ref	Ref	Ref	Ref	Ref
Climate zone H2	0.0188	0.00825	0.00233	0.00935	0.0326*	0.0422*
	(0.0118)	(0.0242)	(0.0134)	(0.0142)	(0.0182)	(0.0234)
Climate zone H1 - coldest	-0.00353	-0.0237	-0.0273	0.0251	0.0284	0.0557
	(0.0224)	(0.0429)	(0.0293)	(0.0313)	(0.0377)	(0.0452)
Town < 2000 inhabitants	0.0614***	-0.0150	0.0114	0.0713***	0.0989***	0.145***
	(0.0205)	(0.0512)	(0.0256)	(0.0233)	(0.0283)	(0.0358)

Town between 2,000 and 10,000 inhabitants	0.0562***	0.0333	0.0239	0.0568**	0.0654**	0.0670*
	(0.0210)	(0.0515)	(0.0248)	(0.0277)	(0.0313)	(0.0391)
Town between 10,000 and 50,000 inhabitants	0.000146	-0.0279	-0.0214	0.00264	0.00990	0.0354
	(0.0207)	(0.0508)	(0.0254)	(0.0226)	(0.0332)	(0.0346)
Town between 50,000 and 200,000 inhabitants	-0.0563**	-0.0413	- 0.0732***	-0.0549**	-0.0669**	-0.0352
	(0.0225)	(0.0563)	(0.0240)	(0.0259)	(0.0326)	(0.0493)
City between 200,000 and 2,000,000 inhabitants	-0.0356*	-0.0467	-0.0445*	-0.0156	-0.0250	-0.0190
	(0.0189)	(0.0499)	(0.0242)	(0.0229)	(0.0270)	(0.0334)
Paris	Ref	Ref	Ref	Ref	Ref	Ref
Constant	0.747***	0.319	0.599***	0.702***	0.732***	1.322***
	(0.121)	(0.249)	(0.148)	(0.141)	(0.199)	(0.213)
Observations	2,044	2,044	2,044	2,044	2,044	2,044
R-squared	0.308					

Note: Robust standard errors in parentheses *** p<0.01, ** p<0.05, ***p<0.1

Table 11: Results of Quantile regression with scale of preferences

	Quantile regression 2				
VARIABLES	0.05	0.25	0.57	0.75	0.90
Energy price in 2012 (log)	-0.771***	-0.591***	-0.375*	-0.744***	-0.986***
	(0.291)	(0.226)	(0.207)	(0.258)	(0.283)
Income (log)	0.0159	0.0215	0.0332**	0.0463*	0.00961
	(0.0244)	(0.0154)	(0.0165)	(0.0238)	(0.0208)
Number of persons	0.0203**	0.0181***	0.0206***	0.0158*	0.0189*
	(0.00881)	(0.00667)	(0.00726)	(0.00830)	(0.0111)
Age inferior to 44 years old	Ref	Ref	Ref	Ref	Ref
Age between 44 and 55 years	0.0109	0.00655	0.0133	0.00188	-0.0321
	(0.0274)	(0.0194)	(0.0196)	(0.0215)	(0.0310)
Age between 56 and 66 years	-0.00826	0.0262	0.0333	0.0454*	0.0612*
	(0.0317)	(0.0210)	(0.0207)	(0.0255)	(0.0314)
More than 66 years old	0.000605	0.0506**	0.0436**	0.0125	-0.000213
	(0.0324)	(0.0230)	(0.0217)	(0.0246)	(0.0362)
tenant	0.00224	-0.000269	-0.0169	-0.0176	-0.0514*
	(0.0252)	(0.0170)	(0.0170)	(0.0222)	(0.0278)
High preference for comfort	0.0310	0.0414**	0.0354**	0.0440**	0.0963***
	(0.0261)	(0.0167)	(0.0168)	(0.0204)	(0.0285)
Medium preference for comfort	-0.00265	0.0382**	0.0214	0.00841	0.0376
	(0.0285)	(0.0182)	(0.0160)	(0.0246)	(0.0279)
Low preference for comfort	0.00983	0.00680	-0.00814	-0.0202	-0.00279
	(0.0253)	(0.0185)	(0.0171)	(0.0201)	(0.0268)
No preference for comfort	Ref	Ref	Ref	Ref	Ref
number of days of housing vacancy during heating period (log)	-0.00223	-0.00654	-0.00495	-0.00560	-0.00973
	(0.00808)	(0.00499)	(0.00534)	(0.00529)	(0.00759)
Number of appliances (log)	0.0590***	0.0782***	0.0687***	0.0442**	0.0385
	(0.0221)	(0.0163)	(0.0197)	(0.0210)	(0.0373)
Class AB	0.452***	0.546***	0.624***	0.851***	1.020***
	(0.0764)	(0.0814)	(0.0562)	(0.144)	(0.122)
Class C	0.332***	0.430***	0.445***	0.450***	0.492***
	(0.0442)	(0.0280)	(0.0264)	(0.0298)	(0.0535)
Class D	0.242***	0.275***	0.300***	0.320***	0.377***
	(0.0349)	(0.0243)	(0.0199)	(0.0267)	(0.0375)
Class E	0.152***	0.184***	0.192***	0.215***	0.209***
	(0.0361)	(0.0235)	(0.0186)	(0.0242)	(0.0314)
Class F	0.0568*	0.108***	0.108***	0.103***	0.0960***
	(0.0323)	(0.0276)	(0.0191)	(0.0252)	(0.0344)
Classe G	Ref	Ref	Ref	Ref	Ref
Surface of the dwelling (log)	-0.103***	-0.152***	-0.191***	-0.193***	-0.227***
	(0.0351)	(0.0199)	(0.0185)	(0.0258)	(0.0306)
Climate zone H3	Ref	Ref	Ref	Ref	Ref
Climate zone H2	0.00732	0.00302	0.00661	0.0350*	0.0580**
	(0.0239)	(0.0139)	(0.0147)	(0.0182)	(0.0239)
Climate zone H1 - coldest	-0.0197	-0.0403	0.0139	0.0320	0.0434
	(0.0431)	(0.0301)	(0.0333)	(0.0343)	(0.0448)
Town < 2000 inhabitants	-0.0145	0.0152	0.0865***	0.101***	0.134***
	(0.0509)	(0.0260)	(0.0242)	(0.0286)	(0.0343)
Town between 2,000 and 10,000 inhabitants	0.0177	0.0271	0.0679**	0.0622**	0.0525
	(0.0528)	(0.0254)	(0.0271)	(0.0317)	(0.0414)
Town between 10,000 and 50,000 inhabitants	-0.0430	-0.0111	0.00479	0.00877	0.0283
	(0.0519)	(0.0270)	(0.0238)	(0.0331)	(0.0351)
Town between 50,000 and 200,000 inhabitants	-0.0502	-0.0725***	-0.0527*	-0.0766**	-0.0370
	(0.0570)	(0.0247)	(0.0274)	(0.0321)	(0.0504)

City between 200,000 and 2,000,000 inhabitants	-0.0668	-0.0404*	-0.0169	-0.0225	-0.0346
	(0.0503)	(0.0238)	(0.0228)	(0.0275)	(0.0321)
Paris	Ref	Ref	Ref	Ref	Ref
Constant	0.354	0.574***	0.748***	0.827***	1.496***
	(0.258)	(0.153)	(0.144)	(0.217)	(0.215)
Observations	2,044	2,044	2,044	2,044	2,044
R-squared					

Note: Robust standard errors in parentheses *** p<0.01, ** p<0.05, ***p<0.1

Table 12: Results of Quantile regression with heating temperature

	Quantile regression, 3				
VARIABLES	0.05	0.25	0.57	0.75	0.90
Energy price in 2012 (log)	-0.819***	-0.681***	-0.432**	-0.648***	-0.688**
	(0.291)	(0.220)	(0.214)	(0.249)	(0.300)
Income (log)	0.0151	0.0100	0.0310**	0.0465**	0.0204
	(0.0254)	(0.0167)	(0.0152)	(0.0213)	(0.0227)
Number of persons	0.0203**	0.0214***	0.0187***	0.0177**	0.0151
	(0.00876)	(0.00717)	(0.00684)	(0.00817)	(0.0107)
Inferior to 44 years old	Ref	Ref	Ref	Ref	Ref
Age between 44 and 55 years	0.00182	0.0102	0.00315	0.00142	-0.0478
	(0.0275)	(0.0207)	(0.0176)	(0.0214)	(0.0305)
Age between 56 and 66 years	0.00240	0.0362*	0.0290	0.0491**	0.0286
	(0.0313)	(0.0201)	(0.0188)	(0.0247)	(0.0330)
More than 66 years old	0.00924	0.0468**	0.0231	0.0193	-0.0263
	(0.0329)	(0.0208)	(0.0203)	(0.0238)	(0.0396)
tenant	-0.000892	-0.00612	-0.00630	-0.00867	-0.0508*
	(0.0257)	(0.0177)	(0.0156)	(0.0231)	(0.0295)
number of days of housing vacancy during heating period (log)	-0.00291	-0.00191	-0.00667	-0.00384	-0.00176
	(0.00827)	(0.00478)	(0.00487)	(0.00529)	(0.00781)
Number of appliances (log)	0.0689***	0.0634***	0.0598***	0.0459**	0.0208
	(0.0231)	(0.0151)	(0.0193)	(0.0210)	(0.0350)
Heating temperature	0.00841	0.0218***	0.0202***	0.0236***	0.0150*
	(0.00690)	(0.00439)	(0.00430)	(0.00512)	(0.00812)
Class AB	0.495***	0.548***	0.600***	0.807***	1.024***
	(0.0737)	(0.0731)	(0.0581)	(0.135)	(0.128)
Class C	0.337***	0.419***	0.415***	0.442***	0.488***
	(0.0451)	(0.0281)	(0.0244)	(0.0297)	(0.0587)
Class D	0.245***	0.272***	0.273***	0.311***	0.375***
	(0.0363)	(0.0254)	(0.0201)	(0.0270)	(0.0364)
Class E	0.156***	0.184***	0.172***	0.196***	0.209***
	(0.0374)	(0.0235)	(0.0166)	(0.0249)	(0.0319)
Class F	0.0598*	0.114***	0.0969***	0.0951***	0.0820**
	(0.0333)	(0.0274)	(0.0174)	(0.0256)	(0.0367)
Class G	Ref	Ref	Ref	Ref	Ref
Surface of the dwelling (log)	-0.116***	-0.132***	-0.174***	-0.182***	-0.205***
	(0.0351)	(0.0184)	(0.0195)	(0.0252)	(0.0331)
Climate zone H3	Ref	Ref	Ref	Ref	Ref
Climate zone H2	0.00385	0.0191	0.0224	0.0397**	0.0614**
	(0.0246)	(0.0134)	(0.0141)	(0.0187)	(0.0253)
Climate zone H1 - coldest	-0.0439	-0.0300	0.0250	0.0176	0.0358
	(0.0389)	(0.0296)	(0.0328)	(0.0348)	(0.0529)

Town < 2000 inhabitants	-0.0212	-0.00118	0.0540**	0.101***	0.151***
	(0.0508)	(0.0235)	(0.0214)	(0.0267)	(0.0405)
Town between 2,000 and 10,000 inhabitants	0.0296	0.00716	0.0404	0.0700**	0.0649*
	(0.0505)	(0.0252)	(0.0256)	(0.0321)	(0.0394)
Town between 10,000 and 50,000 inhabitants	-0.0386	-0.0279	-0.00967	0.00752	0.0161
	(0.0506)	(0.0255)	(0.0202)	(0.0311)	(0.0373)
Town between 50,000 and 200,000 inhabitants	-0.0509	-0.0859***	-0.0624**	-0.0606*	-0.0396
	(0.0587)	(0.0238)	(0.0244)	(0.0352)	(0.0521)
City between 200,000 and 2,000,000 inhabitants	-0.0513	-0.0507**	-0.0278	-0.0229	-0.0181
	(0.0507)	(0.0224)	(0.0206)	(0.0259)	(0.0353)
Paris	Ref	Ref	Ref	Ref	Ref
Constant	0.237	0.230	0.376**	0.297	1.049***
	(0.306)	(0.180)	(0.166)	(0.227)	(0.270)
Observations	2,044	2,044	2,044	2,044	2,044

Note: Robust standard errors in parentheses *** $p<0.01$, ** $p<0.05$, ***$p<0.1$

C.3 Endogeneity test – quantile regression

- Tests of endogeneity

Ho: variables are exogenous

Durbin (score) chi2 (2)=1.27927	p = 0.2580
Wu-Hausman F(1,2019)= 1.26441	p = 0.2610

- Tests of overidentifying restrictions:

Sargan (score) chi2(2) = .984582	p = 0.6112
Basmann chi2(2)= .972526	p = 0.6149

Tests are not significant. Instruments are valid. No endogeneity.