TRIMESTRE DE RECHERCHE 2023

ECONOMIE APPLIQUEE

"DO ENERGY EFFICIENCY IMPROVEMENTS EFFECTIVELY REDUCE THE ENERGY BURDEN CARRIED BY HOUSEHOLDS?"

Supervisors:

Sven HEIM

Guillaume WALD

Presented by:

Ayman ZEINEH

Martin COSTE

TABLE OF CONTENTS

1 Introduction

As the concerns about the climate change rise, and as the demand for a sustainable living increases, the interest in energy efficiency interventions has increased. Particularly in the residential sector. The concept of energy burden, defined as the proportion of household income dedicated to energy expenses, is a metric that encapsulates the economic impact of energy consumption on individual and family budgets. As nations strive to transition towards greener and more sustainable energy practices, understanding the effectiveness of energy efficiency improvements in alleviating the financial strain on households becomes imperative.

This project delves into a pertinent question that resonates with both environmental and socio-economic implications: “Do energy efficiency improvements effectively reduce the energy burden carried by households?”.

This project seeks to investigate the relationship between the renovations that are supposed to improve the energy efficiency and the energy burden of the household. The initial intuitive response is yes. However, upon reflection some problems quickly arise. If a household renovates, was that done with an energy consumption reduction aim in mind or was it just to improve the comfort said household? And if the aim was to reduce the energy consumption, was this renovation a behavior among many that the household is adopting or was it the only action that this household has done to reduce the consumption? As we will discover later, these questions will keep arising whenever a model is adopted.

In order to answer the problem, the following structure was adopted:

An overview of the data, their sources, their details and definitions is firstly addressed. In this section we dive deep into every chosen variable and its function in the context of the topic.

The next section is the methodology that was followed to approach the impact of the energy efficiency improvements on the energy burden, all the while stressing over the need for robust econometric techniques to address the potential biases.

This section goes over a proof of endogeneity, the continuous treatment, the instrumental variable introduced, the two-way fixed effect and the bias on the coefficients of the regressions.

The results of the regressions are then presented and analyzed in terms of statistical significance, magnitude and robustness of the estimated effects.

An interpretation of the results follows and it discusses the regression results, going over the observed biases, and heterogeneity in the effects.

2 Data

2.1 AHS data introduction

The data from which the information about the housing units is drawn is from a survey done by The United States Census Bureau named the American Housing Survey or the AHS for short.

The American Housing Survey (AHS) is sponsored by the Department of Housing and Urban Development (HUD) and conducted by the U.S. Census Bureau.

The survey provides up-to-date information about the size, composition, and quality of the nation's housing and measures changes in our housing stock as it ages. The survey also includes questions about:

- the physical condition of homes and neighborhoods
- the costs of financing and maintaining homes
- the characteristics of people who live in these homes.

Describe the raw data form (4 datasets one for each year that have 69k rows and 2.7k columns)

The selected variables of interest are:

- The renovation cost
- The oil gas and electricity bills in $ to compute the energy consumption of the household
- The household head age to account for the variation of the energy consumption with the age
- The number of people
- The annual revenue in dollars of the household to control for tendencies in behavior that change depending on the wealth level of the household
- The year and the id to perform the regressions and to assemble properly the data and the metropole code to get the temperature data from another source.

Some other variables will be constructed form this basic list and will be used in the regression later on.

2.2 Energy

The oil gas and electricity bills are called OILAMT, GASAMT and ELECAMT respectively. They can take one of the following values:

- 0: Unit does not use fuel oil/gas, elec
- 1heter Vacant home: billed separately
- 2: Occupied or vacant home: included in the rent, site rent, condominium fee, other fee, or other utility bill
- 3: Occupied home: provided free of charge
- 4: 832: $4 to $832
- 833: $833 or more
- N or -6 or -9: Not applicable

From the list above, we cannot benefit from the codes 1, 2, 3, -6 and -9 so they are removed.The entries that remain will hold the dollar amount payed that year by the HU (zeros included). These will be summed up later on to obtain the variable energy consumption in $ for each HU for each year.

2.3 Household head age

The household head age named HHAGE can hold 2 types of values:

- 0: 120: 0 to 120 years
- N or -6: Not applicable (which will be removed in later steps)

The energy demand changes throughout the years for a household depending on the age of the habitants. We add the square of the age to better approximate this behavior as the curve resembles a parabola as proven by "Aging in Population and Energy Demand" [8].

2.4 Number of residents

The number of people in the HU is named NUMPEOPLE and can hold the following values:

- 1: 30: 1 to 30 people
- 30: 30 or more people
- N or -6: Not applicable (which will be removed in later steps)

2.5 ID

The id is named CONTROL, it is a unique eight-digit number that is assigned to each HU. This allows for linking between datasets from different years.

2.6 Income

The annual income of the household is named in the AHS as FINCP, and is in $ and has the following codes:

- -99999999: 99999998: -$99,999,999 to $99,999,998 (with negative values indicating debt)
- 99999999: $99,999,999 or more
- N or -6: Not applicable (which will be removed in later steps)

2.7 Metropole code

The metropole code is named OMB13CBSA it contains codes that represent metropolitan areas that sometimes are a collection of 2 or 3 cities. The possible values are tabulated with the city names and the coordinates that were found separately from the internet.

The list of the possible code that can be found in this column are found in the following table.

Omb13cbsa code	Cities	State
12060	Atlanta-Sandy Springs-Roswell	GA
12580	Baltimore-Columbia-Towson	MD
13820	Birmingham-Hoover	AL
14460	Boston-Cambridge-Newton	MA-NH
16980	Chicago-Naperville-Elgin	IL-IN-WI
17140	Cincinnati	OH-KY-IN
17460	Cleveland-Elyria	OH
19100	Dallas-Fort Worth-Arlington	TX
19740	Denver-Aurora-Lakewood	CO
19820	Detroit-Warren-Dearborn	MI
26420	Houston-The Woodlands-Sugar Land	TX
28140	Kansas City	MO-KS
29820	Las Vegas-Henderson-Paradise	NV
31080	Los Angeles-Long Beach-Anaheim	CA
32820	Memphis	TN-MS-AR
33100	Miami-Fort Lauderdale-West Palm Beach	FL
33340	Milwaukee-Waukesha-West Allis	WI
33460	Minneapolis-St. Paul-Bloomington	MN-WI
35380	New Orleans-Metairie	LA
35620	New York-Newark-Jersey City	NY-NJ-PA

36420	Oklahoma City	OK
37980	Philadelphia-Camden-Wilmington	PA-NJ-DE-MD
38060	Phoenix-Mesa-Scottsdale	AZ
38300	Pittsburgh	PA
38900	Portland-Vancouver-Hillsboro	OR-WA
39580	Raleigh	NC
40060	Richmond	VA
40140	Riverside-San Bernardino-Ontario	CA
40380	Rochester	NY
41700	San Antonio-New Braunfels	TX
41860	San Francisco-Oakland-Hayward	CA
41940	San Jose-Sunnyvale-Santa Clara	CA
42660	Seattle-Tacoma-Bellevue	WA
45300	Tampa-St. Petersburg-Clearwater	FL
47900	Washington-Arlington-Alexandria	DC-VA-MD-WV
99998	All other metropolitan areas	
99999	Not in a metropolitan area	

The code 99998 is rejected as its not an accurate description of the city and therefore cannot be used to get the temperature data, and the code 99999 is rejected as it represents areas that are not metropolitan areas that aren't of interest for the study.

2.8 Renovation cost

The jobtype is the column that contains a code representing the nature of the work that was done on the HU, the list of the possible codes is:

- 01: Earthquake damage required extensive repairs to home
- 02: Tornado, hurricane, etc. damage required extensive repairs to home
- 03: Landslide damage required extensive repairs to home
- 04: Lightning or fire damage required extensive repairs to home
- 05: Flood damage required extensive repairs to home
- 06: Other natural disaster damage required extensive repairs to home
- 07: Added bedroom onto home
- 08: Added bathroom onto home
- 09: Added recreation room onto home

- 10: Added kitchen onto home
- 11: Added other inside room onto home
- 12: Bathroom remodeled
- 13: Kitchen remodeled
- 14: Attached garage or carport added to home
- 15: Porch, deck, patio, or terrace added to home
- **16: Added or replaced roof over entire home**
- **17: Added or replaced siding on home**
- 18: Added or replaced doors or windows in home
- 19: Added or replaced chimney, stairs or other exterior addition
- **20: Added or replaced insulation in home**
- 21: Added or replaced internal water pipes in home
- 22: Added or replaced plumbing fixtures in home
- 23: Added or replaced electrical wiring, fuse boxes, or breaker switches in home
- 24: Added or replaced security system in home
- 25: Added or replaced carpeting, flooring, paneling, or ceiling tiles
- 26: Added or replaced central air conditioning
- **27: Added or replaced built-in heating equipment**
- 28: Added or replaced septic tank
- **29: Added or replaced water heater**
- 30: Added or replaced built-in dishwasher or garbage disposal
- 31: Other major improvements or repairs inside home (up to three could be reported)
- 32: Added or replaced driveways or walkways
- 33: Added or replaced fencing or walls
- 34: Added or replaced swimming pool, tennis court, or other recreational structure
- 35: Added or replaced shed, detached garage, or other building
- 36: Added or replaced landscaping or sprinkler system
- 37: Other major improvements or repairs to lot or yard (up to three could be reported)

Investigating the codes above and returning to the original topic of the study, we select only the codes that can be considered as renovations that improve the energy efficiency of the HU. The list of interesting jobs becomes 16,17,20,27 and 29.

They are selected since these are the renovations governments subsidise in the goal of reducing the energy efficiency. They are therefore expected to reduce the energy consumption by all the governments.

And even if it's not done with the goal of reducing the energy consumption they should be considered since all the governments are funding jobs like that.

This is best illustrated by means of an example:

CONTROL	YEAR	JOBTYPE1	JOBTYPE2	JOBTYPE3	JOBCOST1	JOBCOST2	JOBCOST3	RENOVATION	N	SUM	CUMUL_SUM
11000089	2015							FALSE	FALSE	0	0
11000089	2017	**16**	**17**	01	**800**	**500**	900	**TRUE**	**TRUE**	1300	1300
11000089	2019	30			350			FALSE	**TRUE**	0	1300
11000089	2021	26	**27**		5000	**6000**		**TRUE**	**TRUE**	6000	7300
11000857	2015							FALSE	FALSE	0	0
11000857	2017	**20**	22	25	**2600**	600	5000	**TRUE**	**TRUE**	2600	2600
11000857	2019	**29**	36		**2500**	1900		**TRUE**	**TRUE**	2500	5100
11000857	2021							FALSE	**TRUE**	0	5100

A new variable named R for renovation is created in each dataset i.e. for each year. This variable takes the value true whenever the HU has one or more jobtype columns that contain a job that is in the list of interest.

And from these four R variables a new variable N is created for each year, N being another binary renovation variable but this time once N becomes true it stays true even if the R after the t of the renovation is false.

In parallel to that, the job costs values are also treated. They on the other hand have these possible values:

- 0: 999997: $0 to $999,997
- 999998: $999,998 or more

A new variable is created for each year named sum. This variable holds the total of the job costs spent on the HU on renovation jobs. So only the job types that are in the list of interests are considered and summed in this variable. If a HU does not perform any jobs to enhance the thermal efficiency the sum stays zero.

Then the cumulative sum of job costs is computed and tabulated for each year.

2.9 Intermediate step

After the purging of all these error codes and codes to reject, we find the common ids in the four years studied to obtain the panel data. This is done by using the intersection.

The unit size was supposed to be in the regression but it wasn't stored in the source as a continuous variable. So, we only considered the HUs where the unit size stayed constant over the whole period.

There are now 4 dataframes one for each year that will now be concatenated by rows and the result is the panel data to work on. We also add the HDD columns to these data frames.

2.10 NOAA data

For the HDD, the average monthly temperature data for the US was taken from the National Oceanic and Atmospheric Administration (NOAA) in which the US was mapped onto a grid of squares 5km*5km and that goes from 1895-01-01 to the present time 2023-09-01.

"In these files, monthly maximum, minimum and mean temperature (deg. C. to 100ths) and precipitation (mm to 100ths) are computed for each month from 1895-present. The latest month is appended to these period of record NetCDF files and are based on the GHCN dataset using a 5km gridded approach. The actual grid is 1/24th of a degree. The 1385x596 grid covers CONUS but the non-land points are filled with NaN." - From one of the read me files that accompanies the data files.

The spatial boundaries of the grid are:

West: 172.0

East: -65.0

South: 18.0

North: 72.0

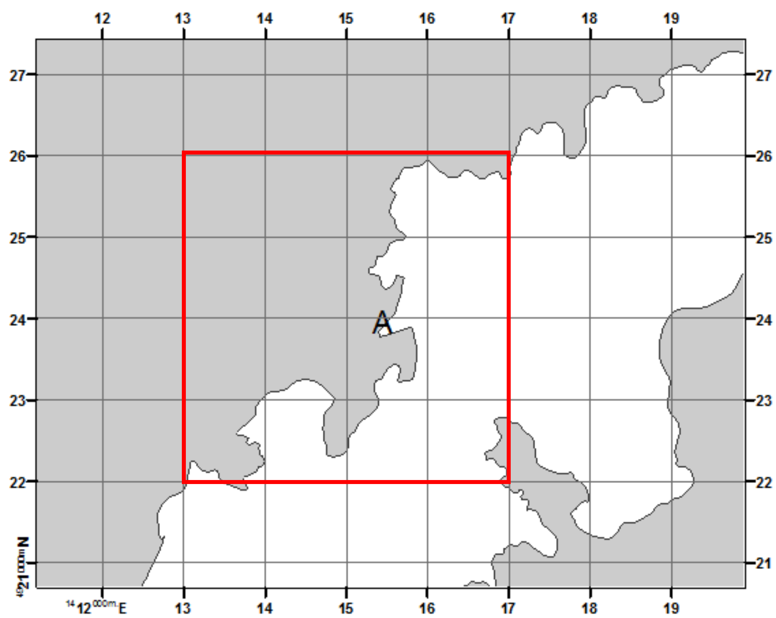
The non-land points have a NaN value in their temperatures.

Two main steps are needed before being able to calculate the HDD of each OMB code:

- Treat the data (reduce the data set to the date period that is relevant to the study 2008 to 2021).
- Approximate the city coordinates onto the us grid. Create the function to approximate the city coordinates to the nearest land point on the grid.

Create the function to locate the nearest point to the OMB code coordinates lat and lon. We take advantage of the fact that the lats are stored in descending order and the lons are stored in ascending order. So, the nearest point on the grid is the last lat that is bigger than the city coordinate, and the last lon that is smaller than the one of the city.

Then if the found coordinate turns out to be a nan aka a non-land point then it cannot be chosen. For that, a 20km*20 km square around this selected point on the grid and the closest land point form the list of these 25 points to the city coordinates is taken as an approximation of the coordinates.



The list of the cities that are in the OMB code are approximated to the grid in this exact manner and we get the approximated latitude and longitude of each metropole in a table.

As a next step, for each OMB code and for each year from 2008 to 2021, the city monthly average temperature is extracted from the NOAA dataset and then used to compute the HDD of the year. This is simply the difference of the temperature and the selected threshold = 18 °C. Then we sum the products of the differences by the number of days in the month to obtain the total amount of HDD in the year and the result is tabulated next to the corresponding OMB code. (insert the formula here)

$$HDD = \sum_{m=1}^{12} (18 - T_m) \times number\ of\ days\ in\ month\ m$$

Now with the HDD data in hand, the average of the last 5 years is computed for the years in question namely 2015, 2017, 2019 and 2021 (HDD average 2017 is average of 2016 15 14 13).

$$\overline{HDD_t} = mean(HDD_{t-1}, HDD_{t-2}, HDD_{t-3}, HDD_{t-4}, HDD_{t-5})$$

Similarly, the relative HDD is defined as follows.

$$\widetilde{HDD}_t = HDD_t - \overline{HDD_t}$$

All the results are tabulated for each OMB code.

Then the HDD is merged with the HUs df according to the OMB code and we obtain a table with the HDD data alongside all the other variables named before.

2.11 Summary of the data

year	count	count renov	% renov	mean sum $	mean cumul $	sum cumul $	mean fincp $	mean age	mean number of people	mean hdd_ave	mean cons $	sum cons $
2015	11,013	2,454	22	1,174	1,174	12,939,298	102,647	53	2.706	1,615	2,163	23,824,968
2017	11,013	2,126	19	1,134	2,309	25,429,822	108,927	54	2.74	1,537	2,298	25,311,408
2019	11,013	2,101	19	1,258	3,567	39,286,323	112,539	55	2.721	1,519	2,350	25,884,408
2021	11,013	2,230	20	1,446	5,014	55,221,959	118,649	56	2.704	1,499	2,349	25,875,840

3 Methodology

In the process of determining the most suitable model for our econometric analysis, we engaged in a thorough exploration of various methodologies. Among the diverse array of methods considered, Instrumental Variation and Two-Way-Fixed-Effects emerged as particularly noteworthy. However, the intricate nature of our research posed challenges in finding and leveraging a viable instrument within the allocated time frame. Consequently, our focus shifted towards the Two-Way-Fixed-Effects model, which proved to be a robust alternative.

Opting for this approach involved incorporating a time fixed effect and a Housing Unit-specific fixed effect, effectively serving as a geographical fixed effect. This intricate setup was achieved by designating the "CONTROL" variable as a fixed parameter. While deliberating on the model specifications, an alternative avenue surfaced — utilizing the "OMB13CBSA" code as a fixed parameter. Although this alternative could have yielded comparable outcomes, a conscious decision was made to persist with the Housing Unit-specific fixed effect for the sake of model consistency.

Our model is as follows:

$$y_{\{it\}} = \beta R_{\{it\}} + \delta X_{\{it\}} + \mu_i + \eta_t + u_{\{it\}}$$

The fundamental structure of our model is anchored in the following components:

"R": Our regressor, signifying the cumulative job costs variables that encapsulate the investments into specific renovations.

"y": The dependent variable denoting the annual energy cost of the Housing Unit at the time "t."

Control variables encompassing income, the number of residents, average Heating Degree Days (HDD), relative HDD, and the age of the head of the Housing Unit.

Navigating the terrain of fixed effects and quantifying the impact of our regressor, denoted as β, necessitates grappling with the endogeneity problem. In classic econometrics, endogeneity arises from sources such as income and political bias. While we have successfully mitigated the endogeneity stemming from income by controlling for the Housing Unit's income, the absence of information pertaining to the political inclinations of the household introduces a layer of complexity. This becomes particularly salient as our hypothesis posits that environmentally conscious homeowners inherently exhibit a proclivity to consume less energy and engage in more renovations.

To address the potential endogeneity related to political bias, a novel variable was introduced to the error term:

$$u = \lambda w + \nu$$

Here, by definition and in alignment with our hypothesis, $\lambda < 0$ and Cov(w, R) > 0. This conceptual framework acknowledges the assumption that environmentally conscious behavior may influence both our regressor and the error term.

Moving forward, we proceed to estimate the results of our regression, recognizing that the derived estimate, denoted as, may be subject to bias. However, the true relationship is more nuanced:

$$Estimation\ of\ \beta = \frac{Cov(y, R)}{Var(R)}$$

As follows given our assumptions:

$$Cov(y,R) = Cov(\beta R + \delta X + \lambda w, R)$$
$$= \beta Cov(R,R) + \lambda Cov(w,R)$$

Thus:

$$\hat{\beta} = \beta + k, \qquad k < 0, \; i.e. \; \hat{\beta} < \beta$$

By incorporating information from our unobserved variable into our regression results, we acknowledge that the outputted estimate may be undervalued. This nuanced consideration enhances the robustness of our estimation, accounting for potential biases associated with unobserved factors and ensuring a comprehensive exploration of the intricacies inherent in our econometric model.

4 Results

OLS Estimation Summary

Dep. Variable:	Y	R-squared:	0.8356
Estimator:	OLS	Adj. R-squared:	0.8356
No. Observations:	42871	F-statistic:	2.105e+05
Date:	Thu, Nov 23 2023	P-value (F-stat)	0.0000
Time:	01:25:14	Distribution:	chi2(8)
Cov. Estimator:	robust		

Parameter Estimates

	Parameter	Std. Err.	T-stat	P-value	Lower CI	Upper CI
C	0.1006	0.0096	10.461	0.0000	0.0817	0.1194
C_SQUARED	-3.177e-07	7.4e-08	-4.2936	0.0000	-4.628e-07	-1.727e-07
FINCP	0.0174	0.0010	17.609	0.0000	0.0155	0.0193
HHAGE	399.29	8.8808	44.961	0.0000	381.89	416.70
NUMPEOPLE	3627.8	49.690	73.009	0.0000	3530.4	3725.2
HDD_AVE	1.8301	0.0560	32.709	0.0000	1.7205	1.9398
RELATIVE_HDD	-0.4870	0.3381	-1.4404	0.1498	-1.1496	0.1757
HHAGE_SQUARED	-2.8893	0.1087	-26.577	0.0000	-3.1024	-2.6763

The findings of our regression analysis yield results that can be characterized as nothing short of astonishing. Within the realm of statistical significance, all variables under consideration display noteworthy correlations with the dependent variable, except for the variable labeled

"RELATIVE_HDD", which stands out as an exception. It is essential to highlight the coherence of the majority of variable coefficients with the underlying model, revealing positive correlations between energy spending and income, the number of occupants in a household, and the frequency of "cold days". Additionally, our results reveal a quadratic relationship between consumption and age, adding depth to our understanding of the complex dynamics at play.

However, the most intriguing revelation in our regression model pertains to the relationship between cumulative renovation investments and energy consumption in housing units. The obtained coefficient, indicating no discernible effect of renovations on energy consumption, is a perplexing outcome, particularly considering its statistical significance. Upon closer examination and by incorporating insights from our methodological approach, we assert that the true coefficient may exceed the one obtained. This leads us to deduce a positive correlation between energy consumption and renovations, contrary to the initial coefficient of zero.

A plausible explanation for this counterintuitive result lies in the concept of a strong rebound effect that will be expanded on when interpreting these results.

5 Interpretation

After all the data purging is done we are left with 11013 entries.

The variables that are statistically significant to Y are:

FINCP HHAGE CUMUL Numpeople HDD_AVE

5.1 Heterogeneity

Understanding the impact of energy efficiency renovations on household consumption is a complex task that requires consideration of various factors contributing to the heterogeneity within the sample. While the overarching goal of energy efficiency renovations is to reduce energy consumption, our analysis reveals intriguing patterns that necessitate a nuanced examination of heterogeneity in the context of this study.

5.1.1 Cross-Sectional Heterogeneity

One of the mains aspects of the results of our study is the evidence of cross-sectional heterogeneity among the households undergoing energy efficiency renovations. The intuitive expectation is the reduction of energy consumption after renovations of this type, however an increase is witnessed on all the sample. This variation can be attributed to the diverse socio-economic characteristics, preferences, and behaviors of households within our sample.

Different income levels, family sizes or household head age may respond differently to energy efficiency renovations. For example, households in the lower percentiles in terms of income might be more likely to invest in renovations to increase comfort without necessarily prioritizing energy saving. In a similar fashion, a larger household cold experience an increase in the energy usage due to augmented demand for heating.

5.1.2 Time-Series Heterogeneity

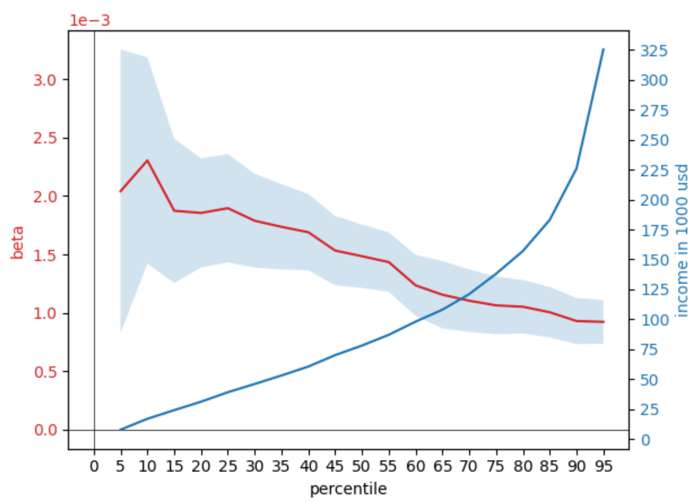
The second dimension of heterogeneity arises when taking into consideration the temporal dynamics of energy efficiency renovations.

Another possible explanation is that there might be an adjustment period right after the renovation during which the households are still adapting to the new technologies or changes in their living environment. This transitional phase may be causing increases in the energy consumption before the long-term benefits of the renovations fully materialize.

5.2 Explanation

This means that there exists a rebound effect, meaning that when a HU performs a renovation to increase the energy efficiency, we expect the beta to be negative, but the result is actually positive. This implies that Hus that renovate increase their energy consumption consistently.

Diving deeper into this last result, effect we divide the HU list into percentiles from 5 to 95 based on the income of the HU. Then the regression with the same variables and settings is done and the resulting betas for each percentile of income as well as their corresponding confidence interval are tabulated and then plotted.



The grayed area is the confidence interval of the beta variable, and its clear that the betas are significant over all the percentiles.

The rebound effect becomes more apparent with the graph as it shows that this effect is actually more extreme in the lower income percentiles. A possible explanation is that in these percentiles, the renovations weren't done with the energy efficiency in mind to reduce the consumption, instead they were done to reduce the needed energy to reach the comfort zone with less power. And when that is achieved the temperature set in these Hus will be raised relative to before the renovation while keeping the bill almost constant. In other words, the kwh would cost them less money and so they increase their energy consumptions.

The only other variable that can help us determine for sure what the real effect of renovations on the energy consumption is the political one i.e. how "green" the HU is. That is because if the household head was "green" they would not only be installing renovations to improve the energy efficiency but they also wouldn't spend as much on energy for environmental purposes.

6 Conclusion

To sum up, the AHS data was first purged from all the entries that cannot be used and a panel data was established. From this data, some selected variables were deemed relevant to the study and were included in the regressions (income, consumption, age, renovation investment, number of people in household). These variables are then used to construct some other variables. The HDD data is extracted from the NOAA data sets and merged with the rest of the variables according to the metropolitan areas.

We implemented a two way fixed effect model to find the relationship between all the control variables and the regressor to the regressand.

In the bias section we find out that the observed coefficient is actually smaller than the theoretical coefficient when the political variable is considered in the regression.

The rebound effect turns out not only to be there but is more prominent in the lower income percentiles. This might be the case because in these percentiles the reduction of the energy consumption is not the primary goal, instead, the renovations are done to improve the comfort for the household.

A notable improvement for future studies is the control for political affiliation, how "green" or "ecological" the household is. This would allow deeper and more accurate results. Another limitation of this study is that fact it wasn't possible to utilize the household size due to the way it is tabulated. An instrumental variable analysis could be beneficial to correctly model the cumulative investment. Another good idea is to try this same methodology of analysis on other countries data to check whether or not the obtained results are exclusive to the US.

7 References

[1] Meredith Fowlie, Michael Greenstone, and Catherine Wolfram. “Do Energy Efficiency Investments Deliver? Evidence from the Weatherization Assistance Program”. In: The Quarterly Journal of Economics 133.3 (2018), pp. 1597–1644. url: https://academic.oup.com/qje/article-abstract/133/3/1597/4828342.

[2] Louis-Gaetan Giraudet, Sebastien Houde, and Joseph Maher. “Moral Hazard and the Energy Efficiency Gap: Theory and Evidence”. In: Journal of the Association of Environmental and Resource Economists 5.4 (2018), pp. 755–790. url: https://EconPapers.repec.org/RePEc:ucp:jaerec:doi:10.1086/698446.

[3] Gael Blaise and Matthieu Glachant. “Quel est l’impact des travaux de renovation energetique des logements sur la consommation d’energie”. In: La Revue de l’energie 646 (2019), pp. 46–60.

[4] Clement de Chaisemartin and Xavier D’Haultfœuille. “Two-way fixed effects and differences-in-differences with heterogeneous treatment effects: a survey”. In: The Econometrics Journal 26.3 (June 2022), pp. C1–C30. issn: 1368-4221. doi: 10.1093/ectj/utac017. eprint: https://academic.oup.com/ectj/article-pdf/26/3/C1/51707976/utac017.pdf. url: https://doi.org/10.1093/ectj/utac017.

[5] Cristina Penasco and Laura Dıaz Anadon. “Assessing the effectiveness of energy efficiency measures in the residential sector gas consumption through dynamic treatment effects: Evidence from England and Wales”. In: Energy Economics 117 (2023), p. 106435.

[6] Kirill Borusyak, Xavier Jaravel, and Jann Spiess. “Revisiting event study designs: Robust and efficient estimation”. In: arXiv preprint arXiv:2108.12419 (2021).

[7] Joshua D Angrist, Guido W Imbens, and Donald B Rubin. “Identification of causal effects using instrumental variables”. In: Journal of the American statistical Association 91.434 (1996), pp. 444–455.

[8] Jaehong Kim and Byeongseon Seo: "Aging in Population and Energy Demand", Korea University, https://eneken.ieej.or.jp/3rd_IAEE_Asia/pdf/paper/066p.pdf