MÁSTER EN INGENIERÍA INDUSTRIAL

TRABAJO FIN DE MÁSTER

CHARACTERISATION OF ENERGY POVERTY IN THREE EUROPEAN COUNTRIES THROUGH THE IMPLEMENTATION OF AN EU MONITORING TOOL

Author: Clément Baumann

Director: José Carlos Romero Mora

Co-Director: Roberto Barrella

Madrid

July 2024

Declaro, bajo mi responsabilidad, que el Proyecto presentado con el título

Characterisation of energy poverty in three European countries through the implementation of an EU monitoring tool

en la ETS de Ingeniería - ICAI de la Universidad Pontificia Comillas en el

curso académico 2023/24 es de mi autoría, original e inédito y

no ha sido presentado con anterioridad a otros efectos.

El Proyecto no es plagio de otro, ni total ni parcialmente y la información que ha sido

tomada de otros documentos está debidamente referenciada.

Fdo.: Clément Baumann Fecha: 11/07/2024

Autorizada la entrega del proyecto

DIRECTORES DEL PROYECTO

José Carlos Romero Mora

Firmado digitalmente por José Carlos Romero Mora
Fecha: 2024.07.16 08:33:01 +02'00'

Firmado por ****2135* ROBERTO BARRELLA el día 15/07/2024 con un

Fdo.: José Carlos Romero Mora Fecha: 15/07/2024
Roberto Barrella

Acknowledgments

I would like to thank all the people who have supported and helped me over the last few months to complete this project.

First, I would like to thank Universidad Pontificia Comillas for the high-quality education and support I have received over the past two years. I want to give a special mention to my thesis directors, José Carlos and Roberto, who accompanied me throughout my work. It has been a great opportunity and a great pleasure to collaborate with them.

Finally, I would like to thank my family and friends for their support during this period.

CARACTERIZACIÓN DE LA POBREZA ENERGÉTICA EN TRES PAÍSES EUROPEOS MEDIANTE LA APLICACIÓN DE UNA HERRAMIENTA EUROPEA DE SEGUIMIENTO

Autor: Baumann, Clément.

Director: Romero Mora, José Carlos.

Co-director: Barrella, Roberto.

Entidad Colaboradora: Catedra de Energía y Pobreza, ICAI – Universidad Pontificia Comillas.

RESUMEN DEL PROYECTO

Este proyecto de investigación presenta una cuantificación de la pobreza energética en España, Italia y Francia, utilizando cinco indicadores diferentes, y un análisis comparativo de la situación de la pobreza energética en estos tres países europeos entre 2017 y 2022. También propone un análisis de los factores de vulnerabilidad a la pobreza energética en Francia.

Palabras claves: Pobreza Energética, Indicadores, Hogares vulnerables, España, Francia, Italia.

1. Introducción

La pobreza energética se refiere a la incapacidad de un hogar de cubrir sus necesidades energéticas básicas de forma adecuada y asequible. Suele ser el resultado de una combinación de limitaciones financieras, viviendas ineficientes desde el punto de vista energético, precios elevados de la energía y factores socioeconómicos desfavorables.

La crisis sanitaria mundial de la COVID-19 y, más recientemente, la crisis energética han agravado considerablemente este problema en la UE. Por ejemplo, la proporción de hogares europeos que no pueden mantener sus casas adecuadamente calientes creció del 6.9% en 2019 al 10.6% en 2023. Estas crisis han destacado la fragilidad de los sistemas energéticos de la UE y han urgido a los responsables de la toma de decisiones a adoptar un enfoque más proactivo para mitigar sus efectos socioeconómicos.

En este contexto, este trabajo está motivado por la creciente urgencia de comprender la pobreza energética en Europa y la intención de contribuir a su medición y a la identificación de los hogares vulnerables en los diferentes países europeos. Este trabajo es necesario para trazar políticas eficaces que aborden adecuadamente este problema.

Para ello, este estudio proporciona una medición de las diferentes dimensiones de la pobreza energética en tres países europeos diferentes, a saber, España, Italia y Francia, así como un análisis de las características de los hogares franceses vulnerables a la pobreza energética.

2. Definición del proyecto

Para medir adecuadamente la pobreza energética, que es un problema multidimensional, en España, Italia y Francia, este proyecto pretende desarrollar una herramienta de medición que calcule varios indicadores para capturar diferentes dimensiones de la pobreza energética. Se han utilizado los cuatro indicadores principales identificados por el Energy Poverty Advisory Hub (EPAH): la proporción de la población que tiene retrasos en pagos de las facturas de servicios públicos, la proporción de la población que no puede mantener su hogar a temperatura adecuada, la proporción de hogares que tiene un gasto energético desproporcionado (2M), y la proporción de hogares que tiene un gasto energético insuficiente (M/2). Además, se ha calculado un quinto indicador: el porcentaje de la población que vive en una vivienda con presencia de goteras, humedades y podredumbre. También se ha realizado un análisis comparativo de la situación de la pobreza energética entre 2017 y 2022 en los tres países.

Además, para identificar las características de los hogares más vulnerables a la pobreza energética, se ha propuesto un análisis de vulnerabilidad que se ha aplicado al caso francés.

3. Descripción del modelo/sistema/herramienta

La herramienta de medición se basa en los datos de las encuestas: Statistics on Income and Living Conditions (SILC) y Household Budget Survey (HBS). Estas dos encuestas, realizadas a nivel nacional de forma periódica, proporcionan información sobre la configuración de los hogares, sus viviendas, su zona de residencia, sus gastos e ingresos y otros datos de interés. Con estos datos se pueden calcular los distintos indicadores de pobreza energética. También se han introducido segregaciones en función del nivel de ingresos y la región de residencia. La herramienta se compone de varios algoritmos escritos en lenguaje R. Su estructura se presenta en la siguiente figura.

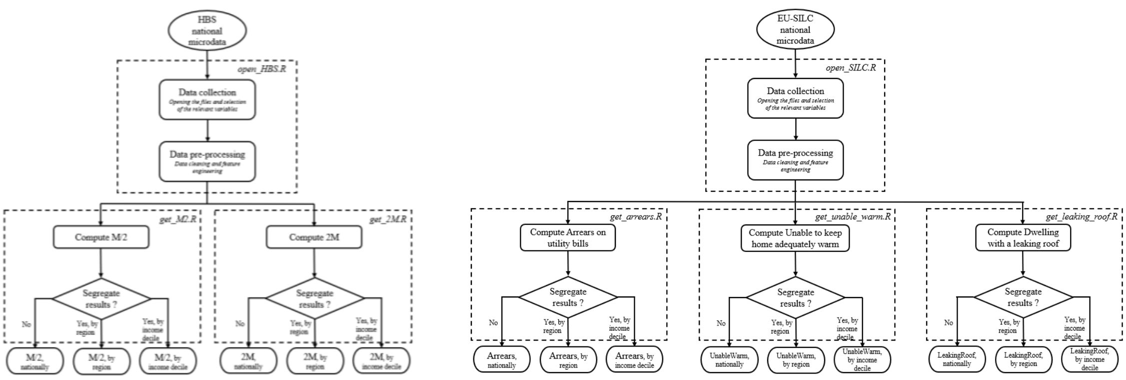

Figura 1. *Esquema funcional de la herramienta de medición*

La metodología utilizada para el análisis de la vulnerabilidad es un análisis econométrico. Los datos proceden de la SILC francesa de 2020. La variable dependiente que se ha utilizado es si un hogar es incapaz de mantener su vivienda a temperatura adecuada. Por lo tanto, el modelo pretende explicar si es probable que un hogar no pueda mantener su vivienda a temperatura adecuada, en función de factores socioeconómicos.

4. Resultados

En España, Italia y Francia, la situación de la pobreza energética es alarmante, ya que los valores de los indicadores han aumentado o se han mantenido estables en el periodo 2017-2022. No obstante, existen diferencias en la prevalencia de los distintos tipos de pobreza energética. El gasto energético desproporcionado es ligeramente superior en España, mientras que el bajo gasto energético absoluto es más frecuente en Francia e Italia. La proporción de la población con atrasos en las facturas de servicios públicos es mayor en España que en Francia e Italia (9.2% frente a 7.1% frente a 5.0% en 2022), pero los tres países comparten una tendencia común con un aumento gradual entre 2018 y 2021 seguido de una pequeña disminución en 2022. Tanto España como Francia muestran resultados alarmantes en la incapacidad de mantener el hogar a una temperatura adecuada en invierno, un indicador que se ha más que duplicado en el período de 5 años, aunque el problema sigue siendo mucho más severo en España. Por el contrario, la proporción de la población italiana incapaz de mantener su hogar adecuadamente caliente disminuyó entre 2019 y 2020, luego se ha mantenido estable en 2021 y ha aumentado en 2022. Además, la proporción de la población que vive en una vivienda con goteras, humedades y podredumbre también ha aumentado alrededor del 50% en los tres países entre 2018 y 2020.

Aunque el modelo econométrico desarrollado no identifica claramente los hogares en situación de riesgo, sí identifica ciertos factores de riesgo. Los resultados del análisis se presentan en la siguiente figura. En pocas palabras, los hogares no propietarios con muchos niños y/o con un solo adulto (monoparentales o monomarentales), que viven en una zona de baja densidad (entorno rural), con un sustentador principal en situación de inestabilidad laboral son los hogares más vulnerables para estar en situación de pobreza energética en Francia.

	Coefficients	Probability ratios
Dummy overpopulated household	-0.0214	0.9788
Type of household		
Single person	0.1047	1.1104
Single-parent family	-0.0464	0.9547
Couple without children	-0.6721**	0.5106**
Couple with at least one child	-0.8105**	0.4446**
Tenure status of households		
Owner	-0.4756*	0.6215*
Tenant or sub-tenant	0.2405	1.2719
Type of house		
Detached house	-0.2693	0.7639
Semi-detached house	0.0368	1.0375
Apartment in a building with less than 10 dwellings	0.1067	1.1126
Apartment in a building with 10 or more dwellings	-0.2853	0.7518
Type of employment of the main breadwinner		
Farmers	-0.1778	0.8371
Artisans, shopkeepers and company managers	0.4802	1.6164
Executives and higher intellectual professions	-0.9346**	0.3927**
Intermediate professions	0.0677	1.0700
Employees	0.6066*	1.8342*
Workers	0.2868	1.3322
Retired people	0.1827	1.2005
Employment of the main breadwinner		
Employed	-0.8808**	0.4145**
Student or trainee	-1.3373***	0.2626***
Unemployed	-0.1406	0.8688
Retired	-0.8541***	0.4257***
Education level of the main breadwinner		
Master or equivalent	-0.9696**	0.3792**
Licence or equivalent	-0.5905	0.5541
DUT or equivalent	-0.5612	0.5705
Bac or equivalent	-0.1733	0.8409
CAP-DNB or equivalent	-0.1000	0.9048
No diploma	0.0736	1.0764
Area of residence		
High-density area	-0.1072	0.8983
Low-density area	0.2045**	1.2269**
Number of children in the household	0.1936***	1.2136***
Number of adults in the household	0.0709	1.0735

Pseudo R-squared = 0.1060
Note: Asterisks indicate the level of significance of the parameter, so that:
*** indicates significance at 1%, ** at 5%, * at 10%

Figura 2. *Factores de vulnerabilidad a no poder mantener el hogar a una temperatura adecuada en invierno en Francia (2020)*

5. Conclusiones

A pesar de algunas disparidades, Francia, Italia y España comparten tendencias similares en términos de pobreza energética. Esto refuerza la idea de que unas políticas armonizadas para abordar la pobreza energética a nivel europeo podrían ser pertinentes, en particular en los países mediterráneos.

Los hogares vulnerables a la pobreza energética siguen siendo difíciles de identificar con precisión debido a la multitud de factores socioeconómicos que hay que tener en cuenta y a las múltiples dimensiones de la pobreza energética. Sin embargo, pueden identificarse algunos factores de vulnerabilidad, como un bajo nivel educativo, estar inactivo o desempleado, o tener una familia numerosa que mantener. Estos factores deberían ayudar a diseñar políticas que aborden adecuadamente la pobreza energética atendiendo a los hogares vulnerables.

6. Referencias

Antepara, I., Papada, L., Gouveia, J., Katsoulakos, N., & Kaliampakos, D. (2020, July 16). Improving Energy Poverty Measurement in Southern European Regions through Equivalization of Modeled Energy Costs . *Sustainability*.

Antunes, M., Teotónico, C., Quintal, C., & Martins, R. (2023). Energy affordability across and within 26 European countries: Insights into the prevalence and depth of problems using microeconomic data. *Energy Economics, 127*.

Baker, K., Mould, R., & Restrick, S. (2018). Rethink fuel poverty as a complex problem. *Nature Energy, 3*(8), 610-612.

Barrella, R., & Romero, J. (2023). Unveiling Hidden Energy Poverty in a Time of Crisis: A Methodological Approach for National Statistics. *Living with Energy Poverty*, pp. 41-51.

Barrella, R., Linares, J. I., Romero, J. C., Arenas, E., & Centeno, E. (2021, October). Does cash money solve energy poverty? Assessing the impact of household heating allowances in Spain. *Energy Research & Social Science, 80*.

Barrella, R., Mora Rosado, S., & Romero Mora, J. (2024). *BOLETIN DE VULNERABILIDAD N°31: El impacto de la pobreza energética en la vulnerabilidad social de la poblacion atendida por CRE en el contexto de la crisis inflacionaria.* Madrid: Cruz Roja Española.

Barrella, R., Romero, J., Linares, J., Arenas, E., Asín, M., & Centeno, E. (2022, April). The dark side of energy poverty: Who is underconsuming in Spain and why? *Energy Research & Social Science, 86*.

Brabo-Catala, L., Collins, E., & Barton, B. (2024). Key Fuel Poverty Indicators and Variables: A Systemic Literature Review. *Economics of Energy, 13*(1).

Camboni, R., Corsini, A., Miniaci, R., & Valbonesi, P. (2021). Mapping fuel poverty risk at the municipal level. A small-scale analysis of Italian Energy Performance Certificate, census and survey data. *Energy Policy, 150*.

Cameron, A. C., & Trivedi, P. K. (2005). *Microeconomics: Methods and Applications.* Cambridge University Press.

Energy Poverty Advisory Hub. (2023). *National Indicators*. Retrieved June 13, 2024, from Energy Poverty Advisory Hub: https://energy-poverty.ec.europa.eu/observing-energy-poverty/national-indicators_en

EPAH. (2022). *Energy Poverty National Indicators: Insights for a more effective measuring.*

EPAH. (2023). *Energy Poverty National Indicators: Uncovering New Possibilities for Expended Knowledge.*

Eurostat. (2021). *Methodological Guidelines and Description of EU-SILC Target Variables.*

Eurostat. (2023, September 13). *9% of EU population unable to keep home warm in 2022*. Retrieved from Eurostat: https://ec.europa.eu/eurostat/web/products-eurostat-news/w/ddn-20230911-1

INE. (2019). Encuesta de presupuestos familiares - 2019.

INE. (2019). Life Conditions Survey - 2019.

INE. (2020). Encuesta de presupuestos familiares - 2020.

INE. (2020). Life Conditions Survey - 2020.

INE. (2021). Encuesta de presupuestos familiares - 2021.

INE. (2021). Life Conditions Survey - 2021.

INE. (2022). Encuesta de presupuestos familiares - 2022.

INE. (2022). Life Conditions Survey - 2022.

INSEE. (2017). Budget des familles - 2017.

INSEE. (2018). Statistiques sur les ressources et les conditions de vie (SRCV) - 2018.

INSEE. (2019). Statistiques sur les ressources et les conditions de vie (SRCV) - 2019.

INSEE. (2020). Statistiques sur les ressources et les conditions de vie (SRCV) - 2020.

International Energy Agency. (2023). *Italy 2023 Energy Policy Review.*

ISTAT. (2019). Household Budget Survey: public use microdata - 2019.

ISTAT. (2020). Household Budget Survey: public use microdata - 2020.

ISTAT. (2021). Household Budget Survey: public use microdata - 2021.

ISTAT. (2022). Household Budget Survey: public use microdata - 2022.

ISTAT. (2023, September 29). *Microdata Files*. Retrieved from ISTAT: https://www.istat.it/en/analysis-and-products/microdata-files

Koukoufikis, G., & Uihlein, A. (2022). *Energy poverty, transport poverty and living conditions-An analysis of EU data and socioeconomic indicators.*

Legendre, B., & Ricci, O. (2015, May). Measuring fuel poverty in France: Which households are the most fuel vulnerable? *Energy Economics, 49,* pp. 620-628.

Menyhert, B. (2023). Existing chalenges and potential improvements to energy poverty measurement. *EPAH Annual Conference.* Warsow.

Meyer, S., Laurence, H., Bart, D., Middlemiss, L., & Maréchal, K. (2018, June). Capturing the multifaceted nature of energy poverty: Lessons from Belgium. *Energy Research & Social Science*, *40*, pp. 273-283.

Palma, P., Barrella, R., Gouveia, J. P., & Romero, J. C. (2024, October). Comparative analysis of energy poverty definition and measurement in Portugal and Spain. *Utilities Policy, 90*.

Romero, J., Centeno Hernáez, E., & Barrella, R. (2022). *Informe de Indicadores de Pobreza Energética en España 2021.*

Romero, J., Linares, P., & López, X. (2018, April). The policy implications of energy poverty indicators. *Energy Policy, 115*, pp. 98-108.

Romero, J., Linares, P., & López-Otero, X. (2017, July). The policy implications of energy poverty indicators.

Thema, J., & Vondung, F. (2020). EPOV indicator dashboard: Methodology guidebook. *Wuppertal Institut für Klima, Umwelt, Energie GmbH.*

CHARACTERISATION OF ENERGY POVERTY IN THREE EUROPEAN COUNTRIES THROUGH THE IMPLEMENTATION OF AN EU MONITORING TOOL

Author: Baumann, Clément.

Director: Romero Mora, José Carlos.

Co-director: Barrella, Roberto.

Collaborating Entity: Chair of Energy and Poverty, ICAI – Universidad Pontificia Comillas.

ABSTRACT

This research project presents a quantification of energy poverty in Spain, Italy and France, using five different indicators, and a comparative analysis of the energy poverty situation in these three European countries between 2017 and 2022. It also proposes an analysis of the factors of vulnerability to energy poverty in France.

Keywords: Energy Poverty, Indicators, Vulnerable households, Spain, France, Italy.

1. Introduction

Energy poverty refers to a household's inability to meet its basic energy needs in an adequate and affordable manner. It often results from a combination of financial constraints, energy-inefficient housing, high energy prices and unfavorable socio-economic factors.

The Covid-19 global health crisis and, more recently, the energy crisis have significantly exacerbated this issue in the EU. For instance, the share of European households unable to keep their homes adequately warm grew from 6.9% in 2019 to 10.6% in 2023. These crises have highlighted the fragility of the EU's energy systems and urged decision-makers to adopt a more proactive approach to mitigating their socio-economic effects.

In this context, this work is motivated by the growing urgency of understanding energy poverty in Europe and the intention to contribute to its measurement and the identification of the vulnerable households in different European countries. This work is required to draw effective policies to adequately address this issue.

To this end, this study provides a measurement of the different dimensions of energy poverty in three different European countries, namely Spain, Italy and France, as well as an analysis of the characteristics of French households vulnerable to energy poverty.

2. Project definition

To adequately measure energy poverty, which is a multi-dimensional issue, in Spain, Italy and France, this project aims to develop a measurement tool calculating several indicators to capture different dimensions of energy poverty has been developed. The four main indicators identified by the Energy Poverty Advisory Hub (EPAH) have been used: the share of the population having arrears on utility bills, the share of the population unable to keep their home adequately warm, the share of the households having a high share of energy expenditure in income (2M), and the share of the households having a low absolute energy expenditure (M/2). Additionally, a fifth indicator has been computed: the share of the population living in a dwelling with presence of leak, damp and rot. A comparative analysis of the energy poverty situation between 2017 and 2022 in the three countries has also been conducted.

Additionally, to identify the characteristics of those households more vulnerable to energy poverty, a vulnerability analysis has been proposed and has been applied to France.

3. Description of the model/system/tool

The measurement tool is based on the data from the surveys: Statistics on Income and Living Conditions (SILC) and Household Budget Survey (HBS). These two survey, conducted nationally on a regular basis, provide information on the configuration of the households, their dwellings, their area of residence, their expenses and incomes, and other data of interest. With these data, the different energy poverty indicators can be computed. Segregations on income levels and region of residence have also been introduced. The tool is composed of various algorithms written in R language. Its structure is presented in the figure below.

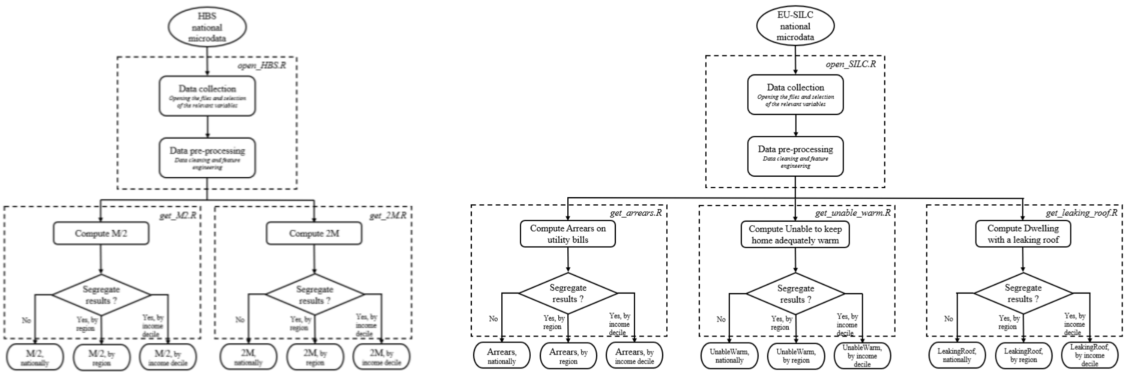

Figure 3. *Functional diagram of the measuring tool*

The methodology used for the vulnerability analysis is an econometric analysis. The data from the 2020 French SILC. The dependent variable that has been used is whether a household is unable to keep its home adequately warm. Therefore, the model intends to explain whether a household is likely to be unable to keep its home adequately warm, based on socio-economic factors.

4. Results

In Spain, Italy and France, the situation of energy poverty is alarming, as the values of the indicators have all increased or remained stable over the 2017-2022 period. Nevertheless, there are differences in the prevalence of different types of energy poverty. Disproportionate energy expenditure is slightly higher in Spain, while low absolute energy expenditure is more prevalent in France and Italy. The share of the population having arrears on utility bills is higher in Spain than in France and Italy (9.2% vs. 7.1% vs. 5.0% in 2022), but the three countries share a common trend with a gradual increase between 2018 and 2021 followed by a small decreased in 2022. Both Spain and France show alarming results in the inability to keep home adequately warm in winter, an indicator that has more than doubled over the 5-year period, although the problem remains much more acute in Spain. Conversely, the share of the Italian population unable to keep their home adequately warm decreased between 2019 and 2020, then it has been stable in 2021 and increased in 2022. Additionally, the share of the population living in a dwelling with leaks, dampness and rot has also increased by about 50% in all three countries between 2018 and 2020.

Although the econometric model developed does not clearly identify households at risk, it does identify certain risk factors. The results of the analysis are presented in the figure below. In a nutshell, non-owner households with many children and/or with one adult (single-parent households), living in a low-density area (rural area), with a main breadwinner in situation of job instability are the most vulnerable households to being in situation of energy poverty in France.

	Coefficients	Probability ratios
Dummy overpopulated household	-0.0214	0.9788
Type of household		
Single person	0.1047	1.1104
Single-parent family	-0.0464	0.9547
Couple without children	-0.6721**	0.5106**
Couple with at least one child	-0.8105**	0.4446**
Tenure status of households		
Owner	-0.4756*	0.6215*
Tenant or sub-tenant	0.2405	1.2719
Type of house		
Detached house	-0.2693	0.7639
Semi-detached house	0.0368	1.0375
Apartment in a building with less than 10 dwellings	0.1067	1.1126
Apartment in a building with 10 or more dwellings	-0.2853	0.7518
Type of employment of the main breadwinner		
Farmers	-0.1778	0.8371
Artisans, shopkeepers and company managers	0.4802	1.6164
Executives and higher intellectual professions	-0.9346**	0.3927**
Intermediate professions	0.0677	1.0700
Office employees	0.6066*	1.8342*
Workers	0.2868	1.3322
Retired people	0.1827	1.2005
Employment of the main breadwinner		
Employed	-0.8808**	0.4145**
Student or trainee	-1.3373***	0.2626***
Unemployed	-0.1406	0.8688
Retired	-0.8541***	0.4257***
Education level of the main breadwinner		
Master or equivalent	-0.9696**	0.3792**
Licence or equivalent	-0.5905	0.5541
DUT or equivalent	-0.5612	0.5705
Bac or equivalent	-0.1733	0.8409
CAP-DNB or equivalent	-0.1000	0.9048
No diploma	0.0736	1.0764
Area of residence		
High-density area	-0.1072	0.8983
Low-density area	0.2045**	1.2269**
Number of children in the household	0.1936***	1.2136***
Number of adults in the household	0.0709	1.0735

Pseudo R-squared = 0.1060
Note: Asterisks indicate the level of significance of the parameter, so that:
*** indicates significance at 1%, ** at 5%, * at 10%

Figure 4. *Vulnerability factors to not being able to keep home adequately warm in winter in France (2020)*

5. Conclusions

Despite some disparities, France, Italy and Spain share similar trends in terms of energy poverty. This reinforces the idea that harmonized policies to tackle energy poverty at the European level could be relevant, especially in Mediterranean countries.

The households vulnerable to energy poverty remain difficult to clearly identify precisely because of the multitude of socioeconomic factors to be considered and the multiple facets of energy poverty. However, some vulnerability factors can be identified, such as a low level of education, being inactive or unemployed, or having a large family to support. These factors should help to design policies to adequately tackle energy poverty by addressing vulnerable households.

6. References

Antepara, I., Papada, L., Gouveia, J., Katsoulakos, N., & Kaliampakos, D. (2020, July 16). Improving Energy Poverty Measurement in Southern European Regions through Equivalization of Modeled Energy Costs . *Sustainability*.

Antunes, M., Teotónico, C., Quintal, C., & Martins, R. (2023). Energy affordability across and within 26 European countries: Insights into the prevalence and depth of problems using microeconomic data. *Energy Economics, 127*.

Baker, K., Mould, R., & Restrick, S. (2018). Rethink fuel poverty as a complex problem. *Nature Energy, 3*(8), 610-612.

Barrella, R., & Romero, J. (2023). Unveiling Hidden Energy Poverty in a Time of Crisis: A Methodological Approach for National Statistics. *Living with Energy Poverty*, pp. 41-51.

Barrella, R., Linares, J. I., Romero, J. C., Arenas, E., & Centeno, E. (2021, October). Does cash money solve energy poverty? Assessing the impact of household heating allowances in Spain. *Energy Research & Social Science, 80*.

Barrella, R., Mora Rosado, S., & Romero Mora, J. (2024). *BOLETIN DE VULNERABILIDAD N°31: El impacto de la pobreza energética en la vulnerabilidad social de la poblacion atendida por CRE en el contexto de la crisis inflacionaria.* Madrid: Cruz Roja Española.

Barrella, R., Romero, J., Linares, J., Arenas, E., Asín, M., & Centeno, E. (2022, April). The dark side of energy poverty: Who is underconsuming in Spain and why? *Energy Research & Social Science, 86*.

Brabo-Catala, L., Collins, E., & Barton, B. (2024). Key Fuel Poverty Indicators and Variables: A Systemic Literature Review. *Economics of Energy, 13*(1).

Camboni, R., Corsini, A., Miniaci, R., & Valbonesi, P. (2021). Mapping fuel poverty risk at the municipal level. A small-scale analysis of Italian Energy Performance Certificate, census and survey data. *Energy Policy, 150*.

Cameron, A. C., & Trivedi, P. K. (2005). *Microeconomics: Methods and Applications.* Cambridge University Press.

Energy Poverty Advisory Hub. (2023). *National Indicators*. Retrieved June 13, 2024, from Energy Poverty Advisory Hub: https://energy-poverty.ec.europa.eu/observing-energy-poverty/national-indicators_en

EPAH. (2022). *Energy Poverty National Indicators: Insights for a more effective measuring.*

EPAH. (2023). *Energy Poverty National Indicators: Uncovering New Possibilities for Expended Knowledge.*

Eurostat. (2021). *Methodological Guidelines and Description of EU-SILC Target Variables.*

Eurostat. (2023, September 13). *9% of EU population unable to keep home warm in 2022*. Retrieved from Eurostat: https://ec.europa.eu/eurostat/web/products-eurostat-news/w/ddn-20230911-1

INE. (2019). Encuesta de presupuestos familiares - 2019.

INE. (2019). Life Conditions Survey - 2019.

INE. (2020). Encuesta de presupuestos familiares - 2020.

INE. (2020). Life Conditions Survey - 2020.

INE. (2021). Encuesta de presupuestos familiares - 2021.

INE. (2021). Life Conditions Survey - 2021.

INE. (2022). Encuesta de presupuestos familiares - 2022.

INE. (2022). Life Conditions Survey - 2022.

INSEE. (2017). Budget des familles - 2017.

INSEE. (2018). Statistiques sur les ressources et les conditions de vie (SRCV) - 2018.

INSEE. (2019). Statistiques sur les ressources et les conditions de vie (SRCV) - 2019.

INSEE. (2020). Statistiques sur les ressources et les conditions de vie (SRCV) - 2020.

International Energy Agency. (2023). *Italy 2023 Energy Policy Review.*

ISTAT. (2019). Household Budget Survey: public use microdata - 2019.

ISTAT. (2020). Household Budget Survey: public use microdata - 2020.

ISTAT. (2021). Household Budget Survey: public use microdata - 2021.

ISTAT. (2022). Household Budget Survey: public use microdata - 2022.

ISTAT. (2023, September 29). *Microdata Files*. Retrieved from ISTAT: https://www.istat.it/en/analysis-and-products/microdata-files

Koukoufikis, G., & Uihlein, A. (2022). *Energy poverty, transport poverty and living conditions-An analysis of EU data and socioeconomic indicators.*

Legendre, B., & Ricci, O. (2015, May). Measuring fuel poverty in France: Which households are the most fuel vulnerable? *Energy Economics, 49,* pp. 620-628.

Menyhert, B. (2023). Existing chalenges and potential improvements to energy poverty measurement. *EPAH Annual Conference.* Warsow.

Meyer, S., Laurence, H., Bart, D., Middlemiss, L., & Maréchal, K. (2018, June). Capturing the multifaceted nature of energy poverty: Lessons from Belgium. *Energy Research & Social Science*, *40*, pp. 273-283.

Palma, P., Barrella, R., Gouveia, J. P., & Romero, J. C. (2024, October). Comparative analysis of energy poverty definition and measurement in Portugal and Spain. *Utilities Policy, 90*.

Romero, J., Centeno Hernáez, E., & Barrella, R. (2022). *Informe de Indicadores de Pobreza Energética en España 2021.*

Romero, J., Linares, P., & López, X. (2018, April). The policy implications of energy poverty indicators. *Energy Policy, 115*, pp. 98-108.

Romero, J., Linares, P., & López-Otero, X. (2017, July). The policy implications of energy poverty indicators.

Thema, J., & Vondung, F. (2020). EPOV indicator dashboard: Methodology guidebook. *Wuppertal Institut für Klima, Umwelt, Energie GmbH.*

Index

Figures Index

I. INTRODUCTION

I.1. Contextualization and justification of the scope

In 2022, 9.3% of the EU population reported not being able to keep home warm during winter (Eurostat, 2023). The growing urgency of understanding and alleviating energy poverty in Europe has highlighted the need for in-depth analysis and targeted action. This master's thesis, entitled "Characterisation of energy poverty in three European countries through the implementation of an EU monitoring tool", seeks to fill this gap by proposing a framework for assessing and monitoring this complex issue in the EU and apply this approach to three case studies: Spain, Italy, and France.

Energy poverty used to refer to issues in home heating, but over the last years, it has evolved towards including all energy needs of a household (Brabo-Catala, Collins, & Barton, 2024). As of now, “Energy poverty is broadly defined as inadequate levels of essential energy services experienced by households. It has various forms and is predicated upon a multitude of vulnerability factors” (Menyhert, 2023). Therefore, energy poverty refers to a household's inability to meet its basic energy needs in an adequate and affordable manner. It often results from a combination of financial constraints, energy-inefficient housing, high energy prices and unfavorable socio-economic factors. In other words, it manifests itself when the energy expenditure of a household compromises its economy, well-being, and safety (Koukoufikis & Uihlein, 2022). Energy poverty may yield profound repercussions within impacted households, precipitating financial stress and deleterious effects on the physical and mental well-being of individuals grappling with this phenomenon (Baker, Mould, & Restrick, 2018).

Due to the various geographical and socio-economic situations in its different countries, there is a great heterogeneity across Europe in energy affordability. Illustratively, the average energy burden (household’s share of energy expenditure over income) varies from 2.9% in Malta to 10.9% in Latvia, the incidence of households experiencing a burden exceeding 10% is minimal in Luxembourg and Malta (2%), contrasting starkly with Hungary, Latvia, and Slovakia, where it exceeds 43% (Antunes, Teotónico, Quintal, & Martins, 2023). These regional disparities underscore the importance of quantifying and monitoring energy poverty at a country level.

Moreover, the Covid-19 global health crisis and, more recently, the energy crisis have significantly exacerbated this problem in Europe. For instance, the share of Spanish households unable to keep home adequately warm grew from 7.6% in 2019 to 14.3% in 2021 (Romero, Centeno Hernáez, & Barrella, 2022) and increased again in 2022 and 2023 reaching values of 17.1% and 20.8% (Energy Poverty Advisory Hub, 2023). These crises have highlighted the fragility of Europe's energy systems and urged decision-makers to adopt a more proactive approach to mitigating the devastating effects of energy poverty.

Against this alarming backdrop, this research aims to develop a novel measurement tool to characterize and monitor energy poverty in Europe and apply this tool to 3 EU Member States, namely Spain, Italy, and France. This tool may help decision-makers better understand regional variations in energy vulnerability and design more effective energy and social policies.

This measurement tool was programmed to read, clean and process the microdata from the Statistics on Income and Living Conditions and Household Budget Survey, which are two surveys conducted at the national level on a regular basis and joined at the EU level by Eurostat. To adequately capture the different dimensions of energy poverty, the tool will allow us to compute five different indicators, namely the share of the population having arrears on utility bills, the share of the population unable to keep their home adequately warm, the share of the population having a high share of energy expenditure in income (2M), the share of the population having a low absolute energy expenditure (M/2), and the share of the population living in a dwelling with presence of leak, damp and rot. It will be built using various algorithms written in R language.

Once energy poverty has been measured, the most vulnerable households need to be identified so that effective measures can be developed to target them and properly tackle this issue. To this extent, this research includes a vulnerability analysis, based on the econometric analysis methodology. Developing a logit model in Python, we identified some key characteristics of the households most likely to be in a situation of energy poverty. This analysis was based on the French Statistics on Income and Living Conditions data from 2020 and used whether a household is unable to keep their home adequately warm as the indicator of energy poverty.

I.2. Objectives and specifications

The thesis aims to address the following specific objectives:

1. Critically **analyze the state of the art on energy poverty measurement.**
2. **Develop a tool to measure energy poverty in European countries** by computing the key energy poverty indicators while considering factors such as income levels, energy expenditure, and auto evaluation reports. The programs developed should allow disaggregating on various socio-economic factors, including income levels or living regions.
3. **Evaluate the impact of socio-economic factors on energy poverty**. Understanding how these factors interact and influence the prevalence of energy poverty is crucial for designing effective mitigation strategies.

As a starting point for the measurement and understanding of energy poverty, we tested the tool and conducted an analysis of energy poverty in three case studies: Spain, Italy, and France.

II. STATE OF THE ART

As highlighted in the introduction, energy poverty is a complex phenomenon, requiring an in-depth understanding of its various dimensions. The origins of household energy poverty encompass a broad spectrum of potential causes. Despite the energy affordability issues mainly affecting poorer households (Antunes, Teotónico, Quintal, & Martins, 2023), having a poor energy performance housing is as important as having a low income in determining energy poverty (Camboni, Corsini, Miniaci, & Valbonesi, 2021). The risk of energy poverty is also increased by the size of the home in relation to the size of the household, and the lower total household expenditures (Camboni, Corsini, Miniaci, & Valbonesi, 2021).

The prevalence of the potential causes of energy poverty varies broadly depending on the country and its geographical and socio-economic characteristics. For instance, in Northern Europe, energy issues consist more of high energy burden and arrears on bills, while in Southern Europe, energy issues consist more of inadequate thermal comfort and under-consumption issues, i.e. the so-called hidden energy poverty (Barrella et al., 2022) (Antepara et al., 2020). Therefore, to adequately identify energy poverty and its different potential causes, employing a variety of pertinent indicators is essential, offering a holistic view of this complex reality, and enabling energy poverty to be quantified and monitored (Antunes et al., 2023).

The EU Energy Poverty Observatory (EPOV), established by the European Union in 2016, tracked and evaluated energy poverty across its member states[1]. A key objective of the EPOV was to propose consistent methodological indicators throughout all member states. For the identification of households at risk of energy poverty, EPOV recognized three primary approaches: comparing household energy expenditure against an absolute or relative threshold (objective or income-expenditure approach), comparing levels of energy service against set standards (direct approach), and analyzing self-assessment reports on their energy situation (subjective approach). Notably, the first two methods essentially rely on quantitative data, while the third relies solely on qualitative data (Thema & Vondung, 2020).

In this line, EPOV introduced four primary indicators:

- Inability to keep home adequately warm: the share of the population that was not able to keep their home adequately warm over the last year.
- Arrears on utility bills: the share of the population that had one or more arrears on utility bills over the last year.
- M/2 (Low absolute energy expenditure): the share of the population which absolute energy expenditure is below half the national median.

[1] In 2021, EPOV was replaced by the Energy Poverty Advisory Hub (EPAH), a new EU project which aims to eradicate energy poverty.

- 2M: the share of the population which share of energy expenditure is above twice the national median (Thema & Vondung, 2020).

On the one hand, the first two primary indicators are consensual-based (subjective), and are computed using self-assessments of the households. On the other hand, the 2M y M/2 are expenditure-based indicators. It has to be noted that these two indicators intend to capture two different dimensions of energy poverty. The 2M aims to quantify energy overconsuming by measuring the share of households having disproportionate energy expenses, while the M/2 attempts to identify hidden energy poverty, meaning that a household is deliberately underconsuming energy in order to save money, by determining the share of households with insufficient energy expenses. However, the latter fails to adequately capture this underconsuming issue (Meyer, Laurence, Bart, Middlemiss, & Maréchal, 2018). For these four indicators, second-level disaggregating variables are considered: income deciles in the country, tenure type, urbanization density, and dwelling type. This aims to identify more precisely the type of population suffering from energy poverty (Barrella & Romero, 2023).

Additionally, the EPOV proposed a battery of secondary indicators. These secondary indicators do not aim to measure energy poverty but to give some insights into the country's situation and help identify eventual causes of energy poverty. They are classified in various categories:

- Energy prices: Fuel oil prices, Biomass prices, Coal Prices, Household electricity prices, District heating prices, Household gas prices.
- Consensual-based: Share of dwelling comfortably cooled in summer time, Share of dwelling comfortably warmed in winter time, Presence of leak, dump, rot.
- Expenditure-based: Share of energy expenditure in income by income quantile.
- Building stock features: Share of dwellings with energy label A, Share of dwellings in intermediately populated areas, Share of dwellings in densely populated areas, Share of dwellings equipped with heating, Share of dwellings equipped with air conditioning, Number of rooms per person by ownership status.
- Poverty and health risks: Poverty risk (AROPE), Excess winter mortality (Thema & Vondung, 2020).

In 2022, the Energy Poverty Advisory Hub (EPAH), the successor of EPOV, revised the primary and secondary indicators. The 24 original secondary indicators were converted and reorganized into 17 secondary indicators (EPAH, 2022). In 2023, EPAH categorized its indicators according to four primary topics and respective subtopics: Climate; Facilities/Housing (subtopics: Building Stock; Energy Consumption; Equipment); Mobility; and Socioeconomic Aspects (subtopics: Socioeconomic and living conditions; Energy expenditure and energy markets; Health) (EPAH, 2023). Figure 1 presents the organization of the indicators among the different topics and subtopics.

Table 1: INDICATORS' TOPICS
(the indicators that cross various themes have a related footnote)

Topic	Subtopic	Indicator
Climate		Cooling degree days
		Heating degree days
Facilities/ housing	Building Stock	Dwellings with energy label A
		Final consumption expenditure of households [1]
		Pop. Liv. Dwelling with presence of leak, damp and rot
		Pop. Liv. Dwelling equipped with heating
		Pop. Liv. Dwelling equipped with air conditioning
		Pop. considering their dwelling as too dark
	Energy Consumption and Equipment	Final consumption expenditure of households [2]
		Final energy consumption in households by energy use
		Final energy consumption in households by type of fuel
Mobility		Final consumption expenditure of households [3]
		Pop. who cannot afford a regular use of public transport
Socioeconomic aspects	Socio Economic and Living Conditions	Arrears on utility bills
		At risk of poverty or social exclusion
		Disposable annual household income
		Inability to keep home adequately warm
		Final consumption expenditure of households [4]
		Housing cost overburden rate
		Pop. Liv. Dwelling comfortably cool during summer time
		Pop. Liv. Dwelling comfortably warm during winter time
	Energy Expenditure and Energy Markets	Energy expenses by income quintile
		Energy Prices
		High share of energy expenditure in income (2M)
		Low absolute energy expenditure (M/2)
	Health	Causes of death
		Excess winter mortality/deaths
		Final consumption expenditure of households [5]
		Pop. Reporting a chronic disease

[1] Indicator's disaggregation: *Maintenance and repair of the dwelling* and *Goods and services for routine household maintenance*.
[2] Indicator's disaggregation: *Water supply and miscellaneous services relating to the dwelling* and *Electricity, gas and other fuels*.
[3] Indicator's disaggregation: *Purchase of vehicles*, *Operation of personal transport equipment* and *Transport services*.
[4] Indicator's disaggregation: *Food and non-alcoholic beverages*, *Actual rentals for housing* and *Imputed rentals for housing*.
[5] Indicator's disaggregation: *Health*.

Figure 5. *Table Presenting the organization of the EPAH indicators in topics and subtopics (EPAH, 2023)*

However, despite these advances, gaps remain in the measurement and characterization of energy poverty in Europe, due to the complexity of the underlying factors and the interactions between economic, social and environmental dimensions (Palma, Barrella, Gouveia, & Romero, 2024). Among the primary indicators, the 2M and M/2 are often criticized, mainly because of the use of relative thresholds. Therefore, the Chair of Energy and Poverty at Universidad Pontificia Comillas has considered the use of two alternative indicators for the Spanish case. First, an indicator based on the minimum income standard (MIS) was developed to offer an alternative disproportionate energy expenditure metric to 2M. According to this indicator, a household will be energy poor if it has an excessive energy expenditure that forces it to do without other basic elements of the basket of needs. It has been considered a relevant indicator since, unlike the 2M indicator, it uses an objective and absolute approximation to obtain the threshold that determines the situation or not of energy poverty due to disproportionate expenditure of a household (Romero, Linares, & López-Otero, 2017). Then, the Chair of Energy and Poverty developed an indicator of hidden energy poverty (HEP). The purpose of this indicator is to go a step beyond under-expenditure and to identify, more precisely, those households that are under-consuming energy for affordability reasons and not for other reasons. Thus, the HEP indicator presents two very relevant innovations with respect to M/2. On the one hand, the expenditure threshold that determines whether the household is underconsuming is obtained through a theoretical energy expenditure model developed by the Chair, namely the RENE model. On the other hand, the indicator calculation incorporates a filter by income deciles that excludes households with higher income (Barrella et al., 2022) (Romero, Centeno Hernáez, & Barrella, 2022).

Criticism can also be directed at subjective indicators. Indeed, they are based on people's lived experiences, which are highly subjective and heterogeneous. Although their consistency across regions and countries is an advantage, there is a clear lack of homogeneity in people’s perceptions across a population. Therefore, objective indicators are necessary to complement subjective indicators (Brabo-Catala, Collins, & Barton, 2024).

Energy poverty indicators are also subject to fluctuations along a calendar year. Therefore, data must be collected and compared in harmonized sampling periods (Menyhert, 2023). In addition, the different indicators give very different poverty rates (Menyhert, 2023). These different rates are explained by the fact that the different indicators make it possible to better identify the different causes of energy poverty, allowing more effective policies to be developed depending on the prevailing problem (Antunes, Teotónico, Quintal, & Martins, 2023). Interestingly, the main indicators of energy poverty have a very low overlap rate, highlighting the existence of different causes of energy poverty, as shown in Figure 2 for the case of Hungary in 2018 and in Figure 3 for the case of Spain in 2015.

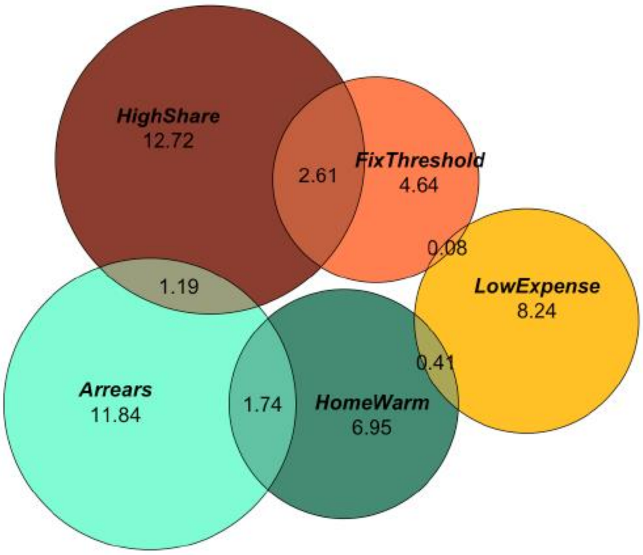

Figure 6. *Five different energy poverty indicators rates and their overlap rates in Hungary in 2018 (Menyhert, 2023)*

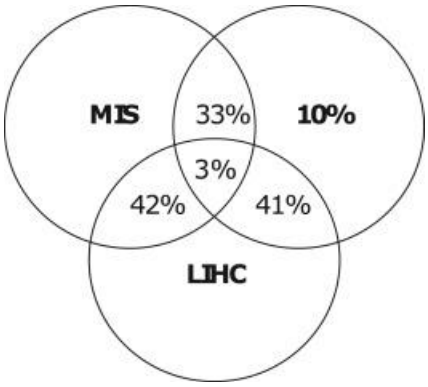

Figure 7. *Three different energy poverty indicators rates and their overlap rates in Spain in 2015 (Romero, Linares, & López, 2018)*

Thus, a combination of subjective and objective variables is relevant because it results in a more holistic view of energy poverty and reduces bias (Brabo-Catala, Collins, & Barton, 2024).

In conclusion, while substantial progress has been made in understanding and measuring energy poverty, ongoing refinements and innovations are necessary to capture the diverse aspects of this multifaceted challenge.

III. METHODOLOGY

III.1. Databases

Two main databases are used to compute energy poverty indicators. On the one hand, the Household Budget Survey (HBS) is a survey conducted at a national level by the national statistics institutes on a periodical basis. It aims to obtain information on the nature and destination of consumption expenditures, as well as on various characteristics related to the socio-economic conditions of the households. This database will be used to calculate objective indicators based on households' income and expenditure data. On the other hand, the European Living Conditions Survey (EU-SILC) is a survey conducted by each national statistics institute and unified by Eurostat, at the EU level. It is carried out every year. The EU-SILC is used to get the subjective indicators. For each of these two surveys, the national sample is made up of around 20,000 responses from the countries studied.

For the Spanish case, both HBS and SILC microdata are publicly available on the *Instituto Nacional de Estadística* (INE) website. The scientific-use microdata of the surveys conducted in 2019, 2020, 2021 and 2022 have been used to measure energy poverty in Spain and its evolution over this period.

For the analysis of energy poverty in Italy, the only database at hand was the HBS public-use data files from the *Instituto Nazionale di Statistica* (ISTAT). The public-use microdata "is developed for some surveys starting from the corresponding scientific-use file to which methods of statistical disclosure control have been applied to reduce the risk of identification of statistical units. In some cases, processing of [public-use microdata files] may produce results that differ from those published or calculated from the corresponding files for research purposes" (ISTAT, 2023). Data from the surveys carried out in 2019, 2020, 2021 and 2022 have been used.

For the study of energy poverty in France, an access to scientific-use files has been granted by the French Data Archives for social sciences. The scientific-use files include confidential household data. It has to be noted that the HBS is not conducted annually in France, but every 4 to 5 years. The data from the 2017 HBS survey and the 2018, 2019, 2020 SILC surveys conducted in France by the *Institut National de la Statistique et des Etudes Economiques* (INSEE) have been analyzed. These datasets are the most recent ones we could have access to.

III.2. Indicators

As highlighted by the energy poverty literature, various indicators have to be used to properly address all the aspects of the problem. For this thesis, the four primary indicators identified by the EPOV were retained for analysis. The (1) Inability to keep its home adequately warm over the last year, the (2) Arrears on utility bills, the (3) M/2, and the (4) 2M indicators have been used (Thema & Vondung, 2020). Additionally, a third consensual-based indicator was considered: the (5) share of the population living in a dwelling with a leaking roof. In this subpart, the methodology used to calculate these energy poverty indicators and the variables used are described in detail.

III.2.a. Inability to keep home adequately warm

The 'Inability to keep home adequately warm' is obtained via a consensual approach based on households' self-assessment. It is obtained directly from the SILC data, specifically from variable HH050 of the EU-SILC. This variable collects the response (Yes or No) to the question to the households about the level of thermal comfort in their home in winter. "The specific question is: Can your household afford to keep its home adequately warm?" (Eurostat, 2021).

This indicator is computed as a percentage of the total population.

III.2.b. Arrears on utility bills

The 'Arrears on utility bills' indicator is also obtained via a consensual approach based on households' self-assessment. As the 'Inability to keep its home adequately warm', this indicator is computed directly from the SILC data, specifically from variable HS021 of the EU-SILC. This variable collects the response to the question to the households about whether or not they experienced any delay in the payment of any utility bill. The specific question is: "In the past twelve months, has the household been in arrears, i.e. has been unable to pay the utility bills (e.g. heating, electricity, gas, water, waste disposal etc.) of the main dwelling on time due to financial difficulties?" (Eurostat, 2021).

This indicator is computed as a percentage of the total population.

III.2.c. Population living in a dwelling with presence of leak, damp and rot

The 'Population living in a dwelling with presence of leak, damp and rot" indicator is also obtained via a consensual approach based on households' self-assessment. As the two indicators previously mentioned, this indicator is computed directly from the SILC data, specifically from variable HH040 of the EU-SILC. This variable collects the response to the question made to the households about whether or not they live in a dwelling with a leaking roof, damp walls, floors or foundation, or rot in window frames of floor (Eurostat, 2021).

This indicator is also computed as a percentage of the total population.

III.2.d. 2M

2M is an objective indicator. It is obtained by comparing household energy expenditure and income against a relative threshold. The data used to compute it comes from the HBS survey.

The 2M is a disproportionate expenditure indicator. It aims to identify those households whose energy expenditure is very high in relation to their income. Specifically, any household whose energy expenditure as a percentage of income is more than twice the national median will be considered energy poor, following the equation:

$$\%EnergyExpenditure_{household} > 2 \cdot Median_{\%EnergyExpenditure_{household}}$$

With: $\%EnergyExpenditure_{household} = \frac{EnergyExpenditure_{household}}{Incomes_{household}}$

The national median used as a reference is computed annually.

The household incomes are calculated from the amount of total net monthly household income of the HBS survey. It is multiplied by 12 to calculate its annual value. To get the income per equivalent person in households of different sizes, the OECD scale of modified factors is applied, using the variable UC2 of the HBS. This variable, known as the OECD-modified scale, takes into account the number of adults (*Nadults*) and the number of children under 14 years old (*Nchildren*) in the household, as shown in the following formula:

$$UC2 = 1 + 0{,}5 * (Nadults - 1) + 0{,}3 * Nchildren$$

For the calculation of household energy expenditures, all the expenditures linked to energy services of the home are considered. These consist of electricity, natural gas, liquefied gas, other liquid fuels, coal and other solid fuels of the main dwelling. For the calculation of the expenditure per equivalent person in households of different sizes, the equivalence factor of the figure 8.

Size of the household (number of members)	Equivalence Factor
1	1,00
2	1,45
3	1,68
4	1,90
5+	1,99

Figure 8. *Equivalence factors table for the calculation of household energy expenditures (Romero, Centeno Hernáez, & Barrella, 2022)*

III.2.e. M/2

M/2 is also an objective indicator. It aims to capture insufficient energy expenditure. The insufficient expenditure indicators consider that a household is energy poor if its energy expenditure is below a certain threshold. The chosen one is M/2, which measures the percentage of households whose energy expenditure is inferior to half the national median, following the equation:

$$EnergyExpenditure_{household} < (Median_{EnergyExpenditure_{household}})/2$$

For the calculation of the household energy expenditure and its national median, the same procedure has been used as for the 2M indicator.

III.2.f. Calculation of the indicators for Italy

As previously explained, the only database we had access to in the case of Italy was the public-use HBS microdata. Thus, we only worked on M/2 and 2M, which are the two EPOV indicators that we can get from the HBS database.

Also, the public-use microdata does not provide all the information the scientific-use microdata does. In that case, the income of the households and the age of the members of the households were not given. Therefore, we used the total expenses of the household as a proxy for its total expenditures, and we used the number of members of the household as a proxy for its UC2. The former is a proxy frequently used in energy poverty literature (Romero, Linares, & López, 2018) (Barrella et al., 2022). To validate the use of the total expenses as a proxy of the total expenditures, we compared the households being under 60% of the median of the proxy with the households being in a relative poverty situation, which are the households which total income is below 60% of median income (index calculated by ISTAT and included in the Italian HBS). There is an accuracy of 83.9%, which validates the proxy (c.f. figure 9, Prediction being the proxy, and Truth the relative poverty situation). These changes only affect the 2M indicator.

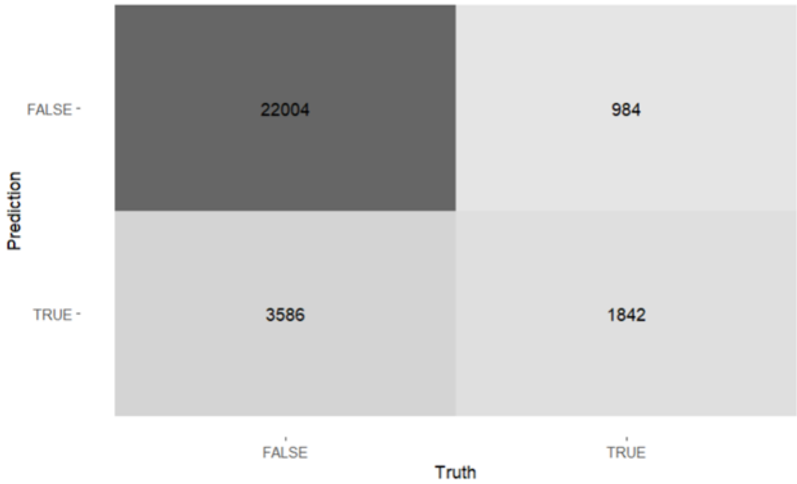

Figure 9. *Correspondence between whether the household is under 60% of the median of the total expenditures (Prediction) and whether the household is in a relative poverty situation (Truth)*

III.3. Implementation and computation of the indicators

The tool developed to read, clean, and process the survey microdata in order to compute the different energy poverty indicators is built using various algorithms written in R language. How it works is shown in the figure 10.

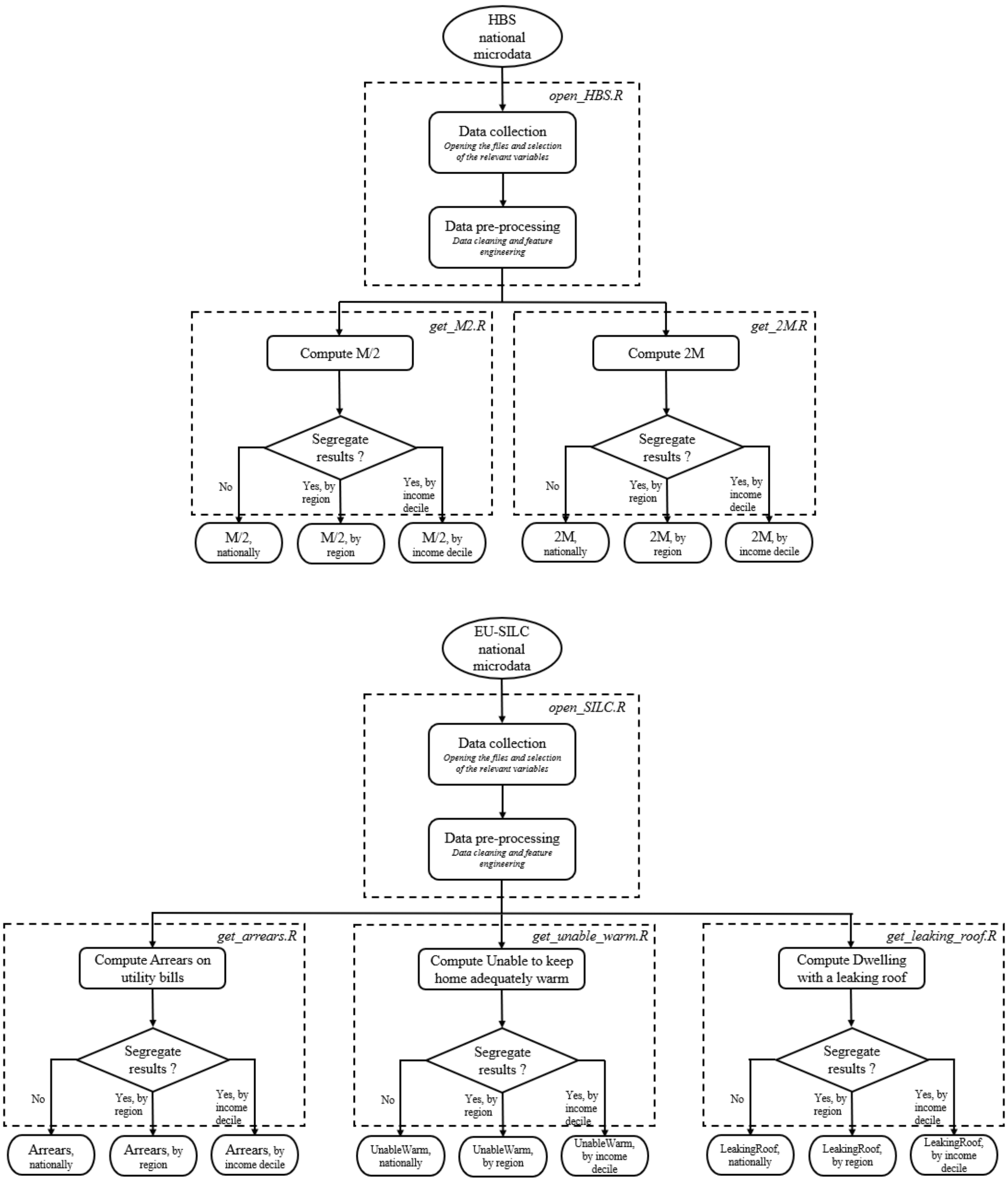

Figure 10. *Functional diagram of the measuring tool*

III.4. Vulnerability analysis

After measuring the different dimensions of energy poverty, finding the characteristics of the households most vulnerable to energy poverty is the next step in drafting effective policies to adequately address this issue. In this manner, policies targeting specific vulnerable groups can be set up to address specific energy poverty dimensions.

The methodology used in this research is an econometric analysis, as proposed by Legendre (Legendre & Ricci, 2015) and also used by Romero (Romero, Linares, & López, 2018). Using Python, a logit model has been developed. The Jupyter Notebook in which the code was developed is available as an appendix to the project. The dependent variable is equal to one if the household is unable to keep their home adequately warm and equal to zero in the other case. The explanatory variables from the 2020 French SILC database that have been used are presented in the figure 11. For each categorical variable, we chose as reference the less probable occurrence.

Variables	**Description**	**Reference**
Dummy overpopulated household	***	***
Type of household	Single person; Single-parent family; Couple without children; Couple with at least one child; Others	Others
Tenure status of households	Owner; Tenant or sub-tenant; Free lodger	Free lodger
Type of house	Detached house; Semi-detached house; Apartment in a building with less than 10 dwellings; Apartment in a building with 10 or more dwellings; Others	Others
Type of employment of the main breadwinner	Farmers; Artisans, shopkeepers and company managers; Executives and higher intellectual professions; Intermediate professions; Office employees; Workers; Retired people; Others	Others
Employment of the main breadwinner	Employed; Student or trainee; Unemployed; Retired; Inactive	Inactive
Education level of the main breadwinner	Doctorate or equivalent; Master or equivalent; Licence or equivalent; DUT or equivalent; Bac or equivalent; CAP-DNB or equivalent; No diploma	Doctorate or equivalent
Area of residence	High-density area; Intermediate-density area; Low-density area	Intermediate-density area
Number of children in the household	***	***
Number of adults in the household	***	***

Source: French SILC in 2020

Figure 11. *Variables in the vulnerability analysis*

According to the logit model methodology, we presume that there exists a latent variable y_i^* such that (1).

$$(1)\ y_i^* = x_i' \cdot \beta + \varepsilon_i$$

where x is the vector of explanatory variables and ε is the error term.

The variable y_i^* is unobservable, we only observe y_i, the binary variable equal to 1 if household i is unable to keep its home adequately warm, and equal to 0 otherwise. It is assumed that the error term follows a standard logistic regression.

Then, the probability of household i to be vulnerable to being unable to keep home adequately warm is given by (2).

(2) $P(y_i = 1) = P(y_i^* > 0) = P(\varepsilon_i > -x_i' \cdot \beta)$

The logit model has the advantage of showing the relationship between the estimated coefficients and the probability ratios. If the probability ratio is superior to one, it means that the probability of being vulnerable to energy poverty is greater and vice versa with respect to the base or reference value of the variable (Cameron & Trivedi, 2005).

As mentioned above, we have used the 2020 French SILC database, which features a sample of 10673 households after depurating the data. All the variables included in the model are dummies, including for each category as many dummies as there are alternatives, except for one, known as the base or reference, on which the probability coefficient is calculated.

IV. RESULTS

IV.1. Measuring energy poverty

This section presents the results of the quantification of energy poverty in Spain, Italy and France (Sections IV.1.a, IV.1.b. and IV.1.c.), then comparing the situation of energy poverty in these countries (Section IV.2).

IV.1.a. Energy poverty in Spain

In the wake of the energy and inflationary crisis that emerged in the post-COVID-19-pandemic era, economic difficulties have surged in Spain, especially affecting the most vulnerable populations in the society (Barrella, Mora Rosado, & Romero Mora, 2024). This subsection presents the results of energy poverty measurement in Spain between 2019 and 2022.

In the case of Spain, we computed the four main EPOV energy poverty indicators, namely the share of the population having arrears on utility bills, the share of the population unable to keep their home adequately warm, the 2M indicator, and the M/2 indicator. The "Population living in a dwelling with presence of leak, damp and rot" indicator was only computed for the years 2019 and 2020, as the EU-SILC survey did not include the question about the presence of leak, damp and rot in the dwelling in 2021 and 2022.

First, we start by analyzing the results at a national scale, which are presented in the figure 12. Globally, between 2019 and 2022, the situation of energy poverty in Spain has worsened, with every indicator being higher in 2022 than in 2019. This situation is alarming.

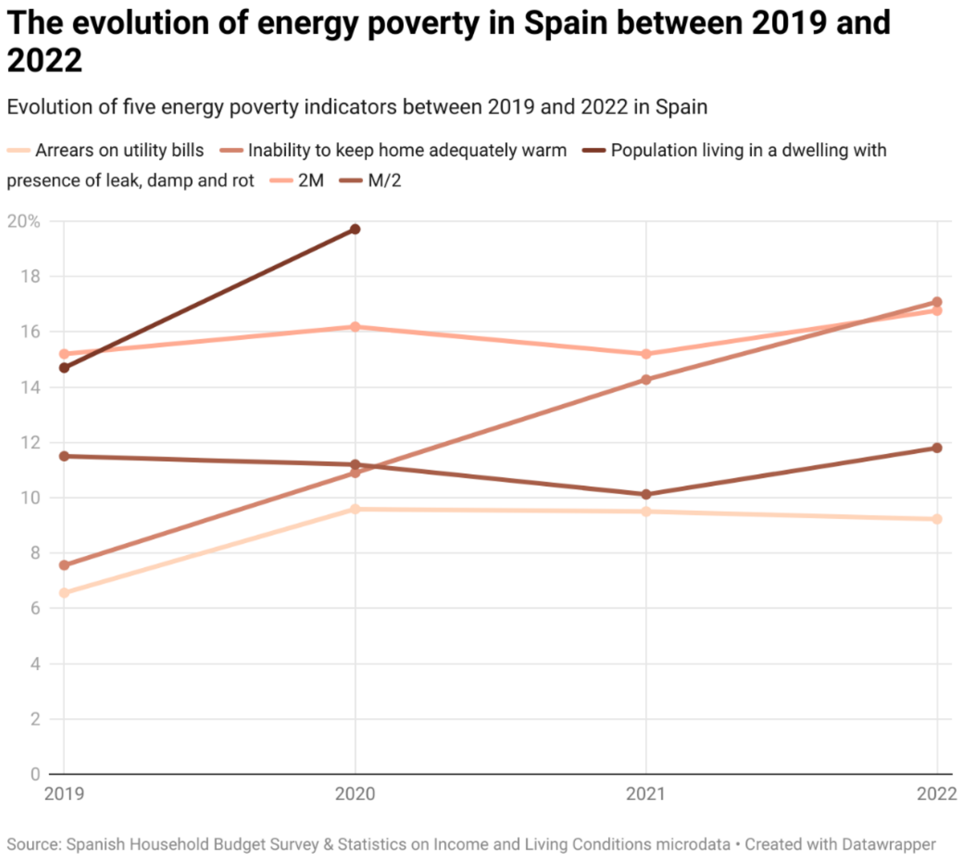

Figure 12. *Evolution of energy poverty indicators in Spain between 2019 and 2022*

After an increase from about 6.5% to 9.5% between 2019 and 2020, the arrears on utility bills indicator remained around 9.5% over the last three-year span. In other words, every year between 2020 and 2022, almost 10% of Spanish citizens reported having arrears in the payment of energy bills. On the one hand, it is observed that the indicator does not show a worsening of payments delay during the period, but at the same time it should be noted that 10% of households with delayed bills is still a very worrying figure. This indicator is still well above its pre-COVID-19 pandemic and pre-energy crisis value, and there has been no sign of improvement.

Then, the indicator of being unable to keep home adequately warm is the one that shows the most worrisome evolution. It surged from 7.56% in 2019 to 17.08% in 2022. This means that more than 17% of citizens declared not to have adequate thermal comfort conditions in winter. A possible explanation for this skyrocketing increase is that households have been controlling more and more their energy expenses, even sacrificing thermal comfort in their homes, due to energy costs and inflation in general.

The share of the Spanish population living in a dwelling with the presence of leak, damp and rot is also of concern, going from 14.70% in 2019 to 19.71% in 2020. This highlights the need to renovate a large number of homes in Spain. The confinement caused by the global pandemic and the consequent decrease in purchasing power has undoubtedly prevented many renovations. In any case, there is an urgent need to speed up the renovation of many residential buildings.

Concerning 2M and M/2 indicators, after a small decrease in 2021, the indicators exceeded their 2020 values in 2022. In 2022, 16.77% of Spanish households spent more than double the national median on energy; and 11.80% of Spanish households had a share of energy expenditures over total income inferior to half the national median.

When segregating the results by total income deciles, the indicators' values decrease when the total income, which is a proxy for wealth, increases. These results were expected, as we can assume that the wealthier a household is, the fewer difficulties it will have in meeting its energy needs (see Figure 13). It has to be noted that among the first two deciles, the disproportionate expenditure issue, which corresponds to the 2M indicator, is prevalent, while in the middle classes, the inability to keep their home adequately warm becomes the most important issue. This highlights the importance of adequately identifying the vulnerable households to each issue in order to tailor the solutions and help to alleviate energy poverty.

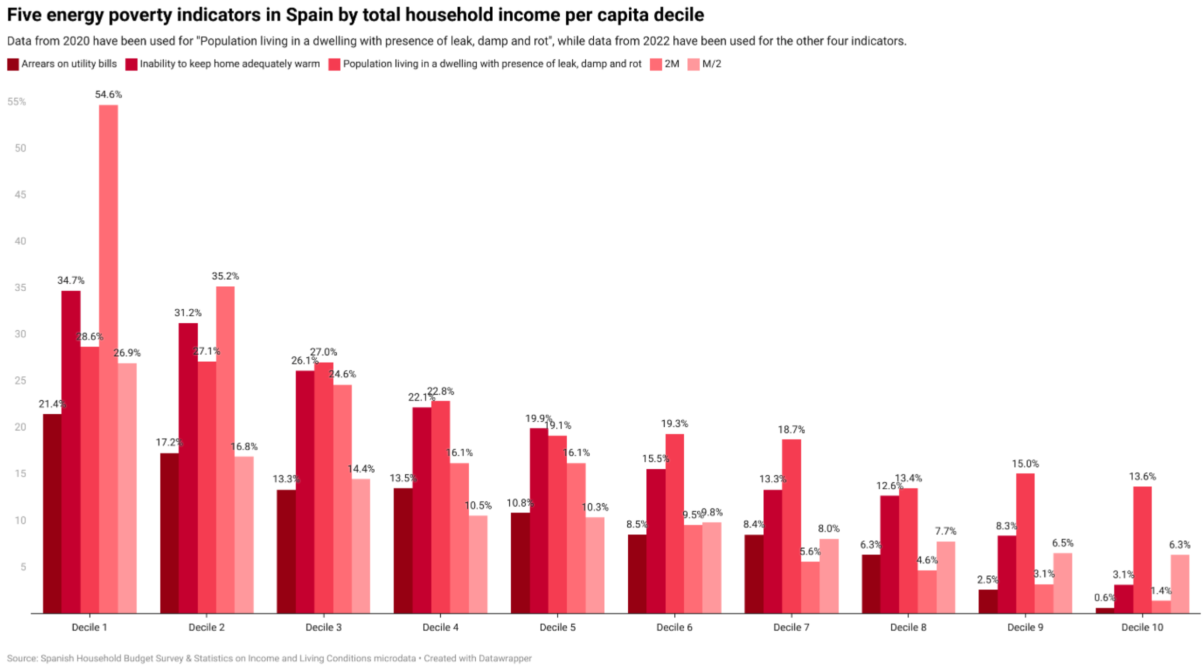

Figure 13. *Values of five energy poverty indicators in Spain, by household income per capita deciles*

Regarding the geographical repartition of energy poverty, figures 14 to 18 presents the values of the five energy poverty indicators in each Spanish region in 2022. Four of the five indicators, the arrears on utility bills, the inability to keep home warm, the population living in a dwelling with presence of leak, and the M/2, show a prevalence of energy poverty in the southern regions of Spain, which are also among the less wealthy in the country. Also, there is a prevalence of disproportionate energy expenditures (2M) in the colder and more rural regions of the country (Barrella, Linares, Romero, Arenas, & Centeno, 2021).

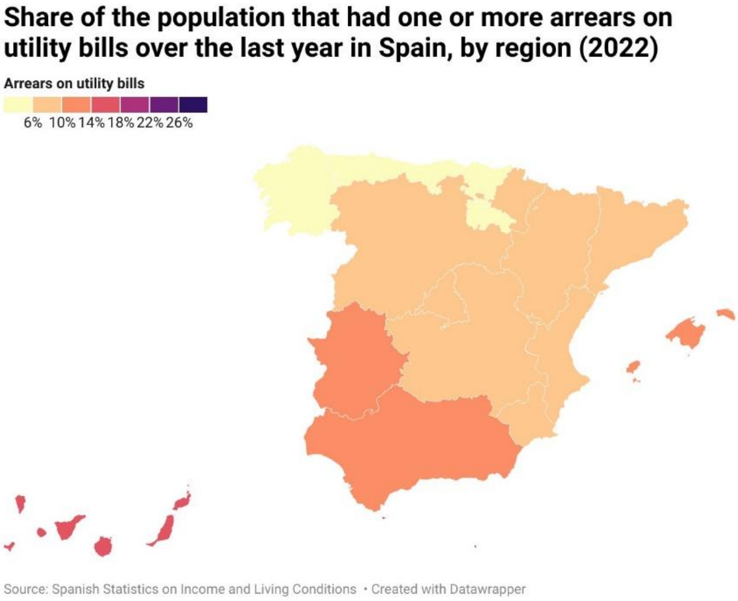

Figure 14. *Share of the population with arrears on utility bills in Spain, by region*

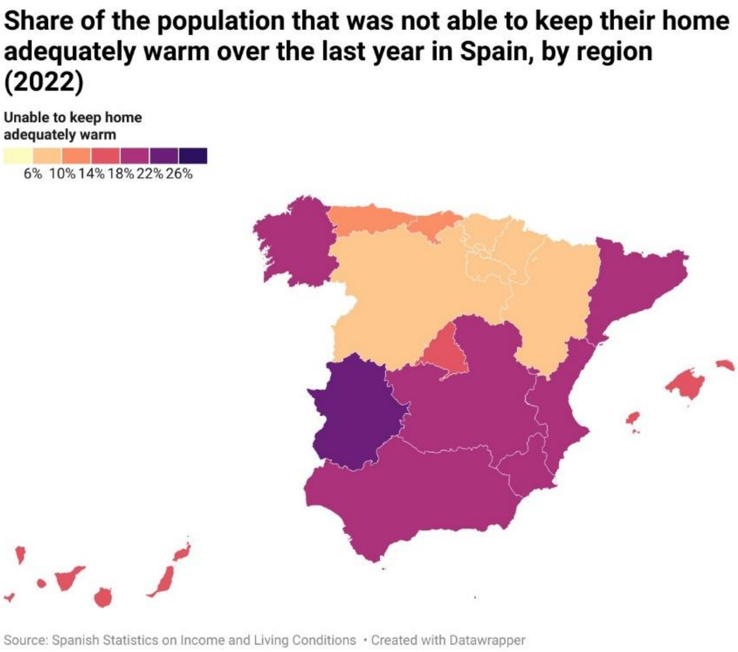

Figure 15. *Share of the population that is not able to keep their home adequately warm in Spain, by region*

The higher inability to keep home adequately warm in the southern regions, which are also the ones with higher temperatures seems contradictory. This can be explained by the type of constructions used there, as pointed out in previous studies (Barrella et al., 2022). Houses and buildings are not built for cold winters. Therefore, they are more difficult to heat in winter, making it difficult to keep them adequately warm in winter. Other variables could also contribute to this result, such as purchase power, income being generally lower in southern regions.

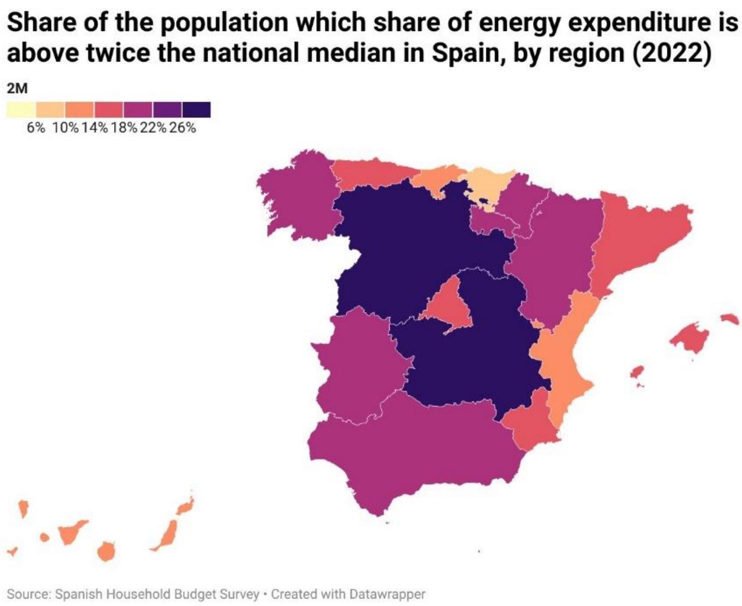

Figure 16. *Share of the Spanish population which share of energy expenditure is above twice the national median, by region*

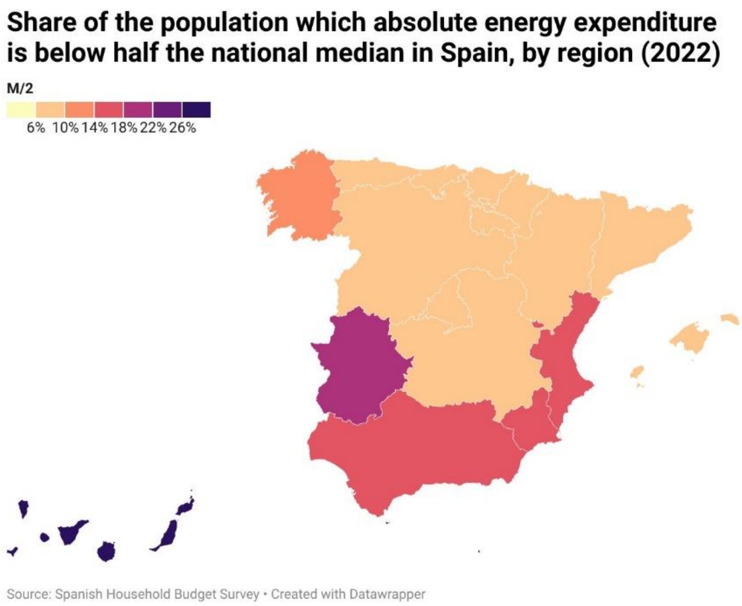

Figure 17. *Share of the Spanish population which absolute energy expenditure is below half the national median, by region*

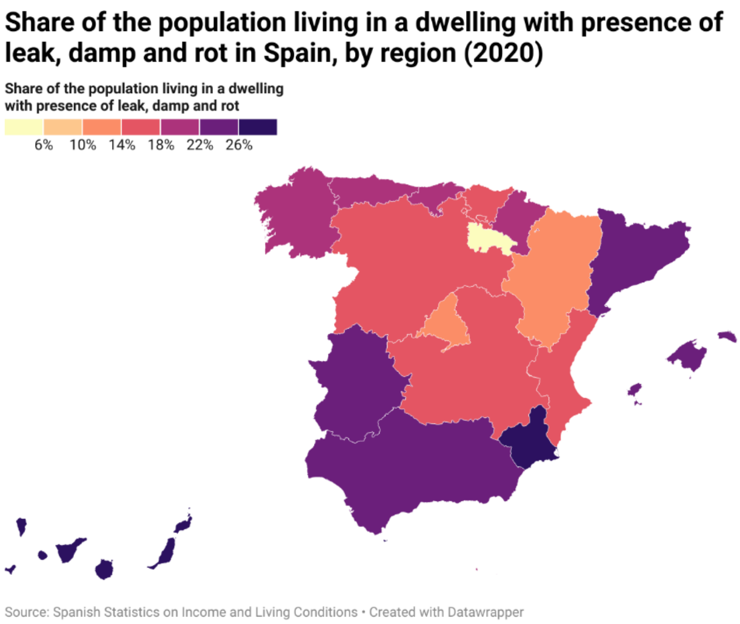

Figure 18. *Share of the Spanish population living a in a dwelling with presence of leak, damp and rot*

IV.1.b. Energy poverty in Italy

In the 2019 Italian National Energy and Climate Plan, the government identified the need to address energy poverty as part of a just energy transition (International Energy Agency, 2023). The objective of this subpart is to see, in light of the available data, what happened in Italy since 2019 in terms of energy poverty. Thus, in this section, we will present the results obtained for the main indicators of energy poverty in Italy from 2019 to 2022.

When analyzing the results obtained (see figure 20), we observe that the 2M indicator remained stable over the period, being around 16%. This means that 16% of Italian households spent more than twice the national median on energy expenses.

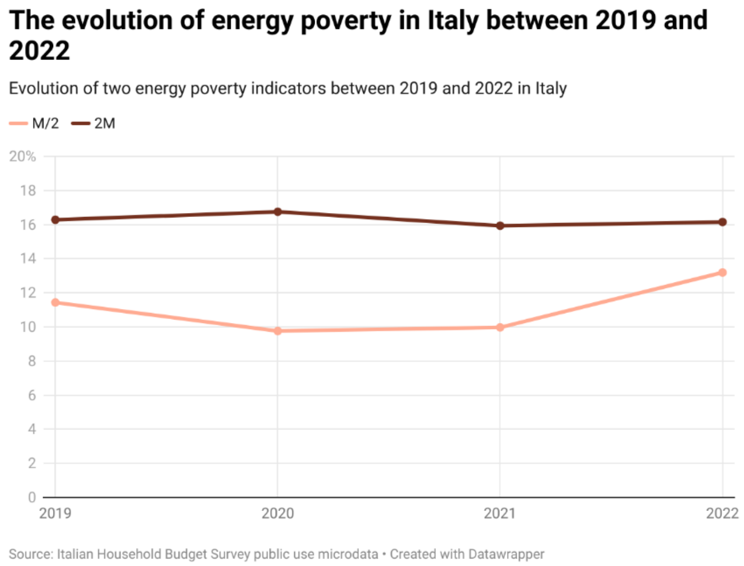

Figure 19. *Evolution of energy poverty indicators in Italy between 2019 and 2022*

Then, the M/2 indicator decreased in 2020 before significantly increasing in 2022. This indicator captures the households that spend in energy per person less than half the national median. In 2022, the national median of energy expenses per person of the households increased to 1081 EUR, from 856 EUR in 2021 (856 EUR in 2019, and 816 EUR in 2020). This increase in the threshold value of the indicator might be the reason behind the significant increase in the M/2 indicator.

When segregating the results by "income" (total expenses per capita) deciles, the indicators' values decrease when the total expenses, which is a proxy to wealthiness, increase. These results were expected, as we can assume that the wealthier a household is, the fewer difficulties it will have to meet its energy needs. Looking at the 2M and M/2 expenses deciles graphs, a quite interesting result is that in 2022 the 2M indicator decreased in the two lowest expenses deciles while increasing in the other ones. In the same year, the M/2 increased in all deciles.

We also note that the share of disproportionate energy expenditure (2M) based on total household expenditure is much higher than the share of energy under-consumption (M/2). Energy poverty reduction policies targeting the poorest Italian households might, therefore, need to focus on the problem of disproportionate energy expenditure. A potential solution to mitigate this problem would be to facilitate the thermal renovation of buildings occupied by the poorest households.

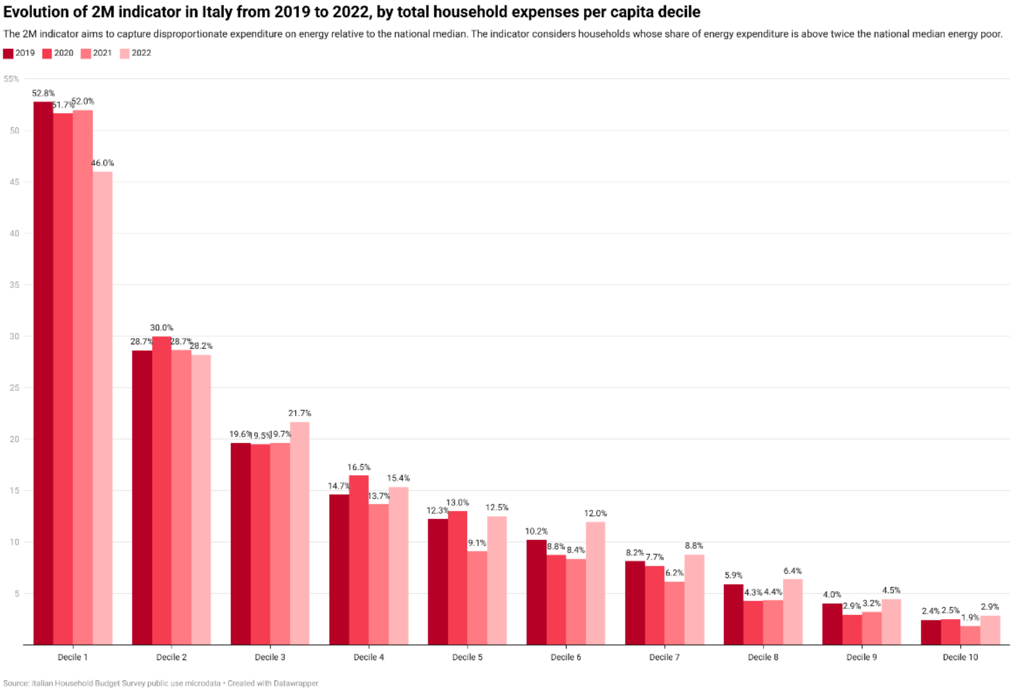

Figure 20. *Evolution of 2M indicator in Italy between 2019 and 2022, by total household expenses per capita decile*

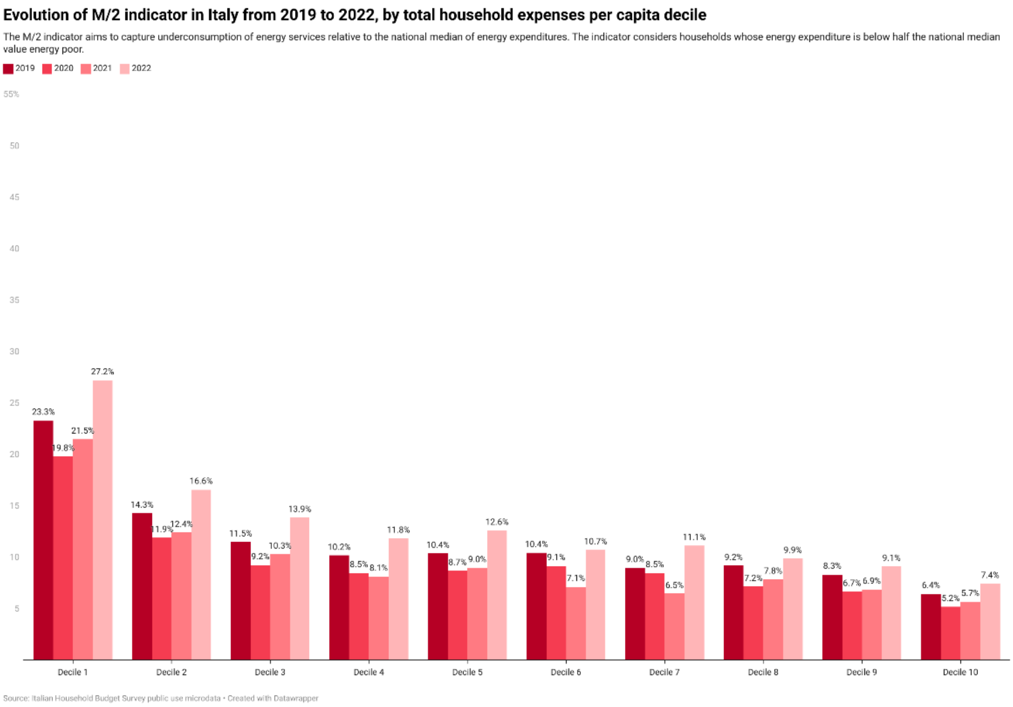

Figure 21. *Evolution of M/2 indicator in Italy between 2019 and 2022, by total household expenses per capita deciles*

The energy poverty situation also differs between the different regions of Italy (see figures 23 and 24).

Disproportionate expenditure on energy is higher in the southern regions of Italy, with Calabria, Basilicata and Molise having the highest 2M value. There seems to be a north–south divide regarding this indicator. There could be several reasons why the share of energy expenditure among the total household expenses might be higher in southern Italy, including income, reliance on less efficient heating systems, poorly insulated homes, and different energy pricing structures in the north and south of Italy.

On the other hand, underconsumption (M/2) seems to be a bit higher in the Northern part of Italy, with Liguria and Trentino among the three regions with a higher M/2 value. The energy expenditure being lower in northern Italy could be explained by more efficient heating systems and better insulated homes in regions that are richer than the southern ones. If that is the case, it is not really energy poverty.

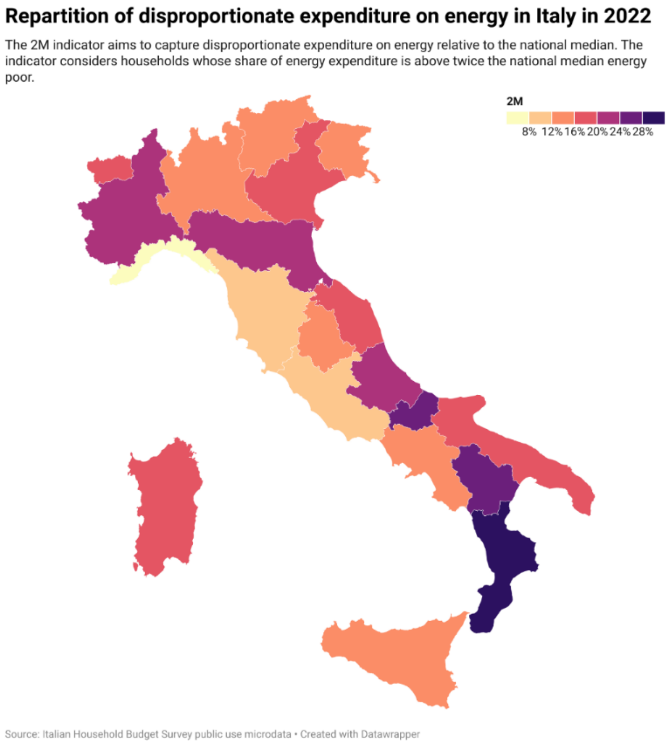

Figure 22. *2M indicator in Italy in 2022, by region*

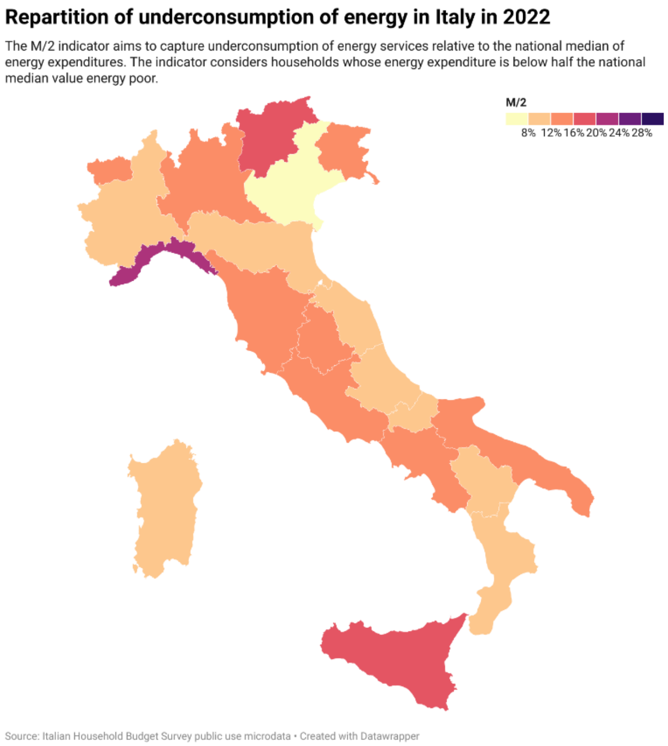

Figure 23. *M/2 indicator in Italy in 2022, by region*

IV.1.c. Energy poverty in France

This subsection presents the results of the measurement of energy poverty in France. For the case of France, the most recent data we had access to are the Household Budget Survey data from 2017 (INSEE, 2017) and the Statistics on Income and Living Standards from 2018 to 2020 (INSEE, 2018) (INSEE, 2019) (INSEE, 2020). Contacts have been made with Réseau Quetelec to gain access to the scientific-use microdata. The national results obtained are presented in the figure 25.

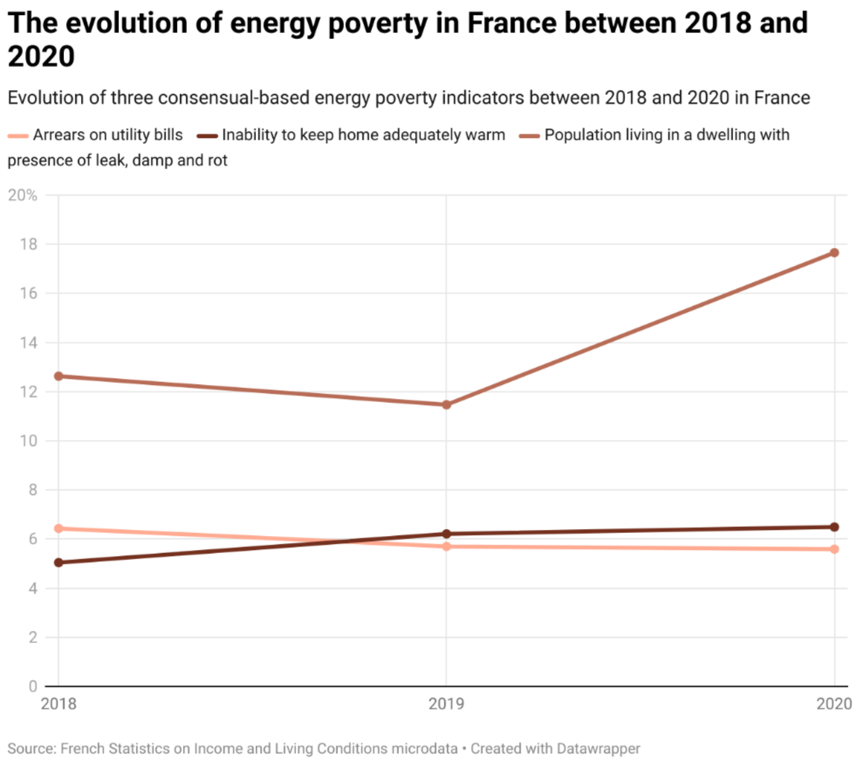

Figure 24. *Subjective energy poverty indicators in France between 2018 and 2020*

First, the situation regarding the arrears on utility bills has slowly improved between 2018 and 2020. In 2020, 5.59% of the French population, which corresponds to more than 3.5 million people, was experiencing arrears on utility bills at least once over the last year. This improving situation is positive, especially knowing that the COVID-19 crisis blew up in 2020 and had negative economic consequences.

Conversely, over the three-year span, the share of the population unable to keep home adequately warm has slightly increased from 5.04% to 6.49%. That is to say that about 4,5 million French citizens were unable to maintain their homes warm enough in 2020 winter. This situation must be closely monitored, and the results from the following years will be interesting due to the energy crisis that arose in 2021. Indeed, according to Eurostat, after a slight decrease to 6%, the Inability to keep home adequately warm indicator grew to 10.9% in 2022 (Energy Poverty Advisory Hub, 2023). The doubling of this indicator's value in 5 years is very alarming and may be explained by households controlling more their energy expenses, even sacrificing thermal comfort in their homes, due to energy costs and inflation in general.

Another worrying result is the evolution of the indicator "Dwelling with presence of leak, damp and rot". From 2019 to 2020, it went from 11.47% to 17.66% of the population living in a dwelling with the presence of leaks, dampness and rot. This increase is probably linked to the COVID crisis. A possible explanation is that the economic difficulties experienced by some of the citizens prevented them from carrying out maintenance work in their homes.

Additionally, in 2017, 15.65% of French households spent more than twice the national median on energy expenses, and 20.76% of French households spent on energy less than half the national median.

When segregating the results by total expenses deciles, the indicators' values decrease when the total expenses, which is a proxy to income, increases. These results were expected, as we can assume that the wealthier a household is, the fewer difficulties it will have to meet its energy needs. For the poorest households, the two most prevalent issues are the disproportionate energy expenses (2M) and the underconsumption (M/2). The prevalence of 2M decreases progressively as the household income per capita increases, and the most prevalent issues for the higher income deciles are underconsumption and the presence of leak, damp and rot in the dwelling. From these results, we can deduce that policies addressing the disproportionate energy expenses, the arrears on utility bills and the inability to keep home adequately warm should target the poorer households, while the ones addressing the presence of leak in the dwelling could be efficient for all the population. Finally, as mentioned in the case of Italy, the M/2 indicator cannot clearly identify households in energy poverty.

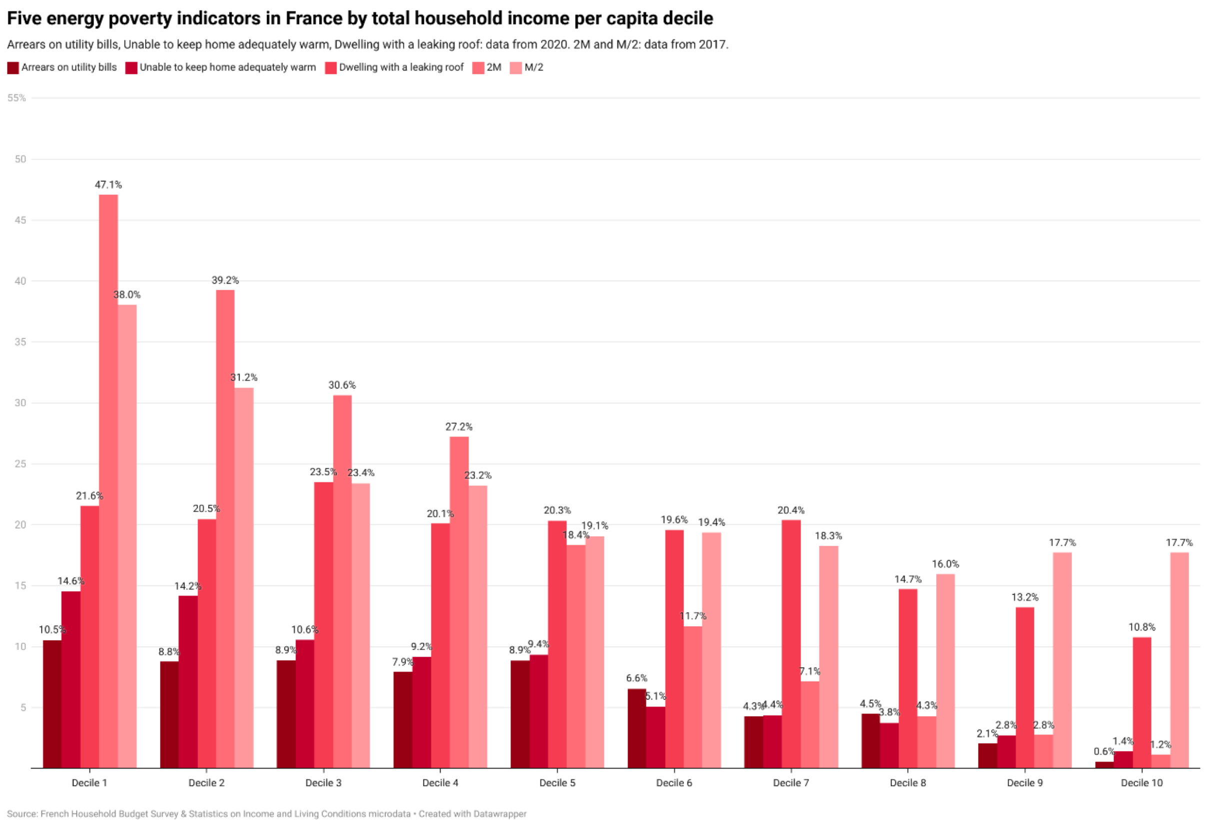

Figure 25. *Values of five energy poverty indicators in France, by total household income per capita deciles*

Overall, there is a relative geographical homogeneity in the values of the various energy poverty indicators in France. However, some variations can be observed. Among the most significant ones, the share of the population living in a dwelling with the presence of leaks in the Paris Region is higher than average, which could be explained by the poverty of some suburbs of Paris. Also, in the Mediterranean Region a higher share of the population is experiencing difficulties in keeping their home adequately warm, despite being one of the warmer regions of France. As for the Spanish case, this might be explained by the type of constructions and heating equipment used there. Dwellings are usually less efficient in the south of France. Thus, they are more difficult to heat in winter, making it difficult to keep them adequately warm in winter.

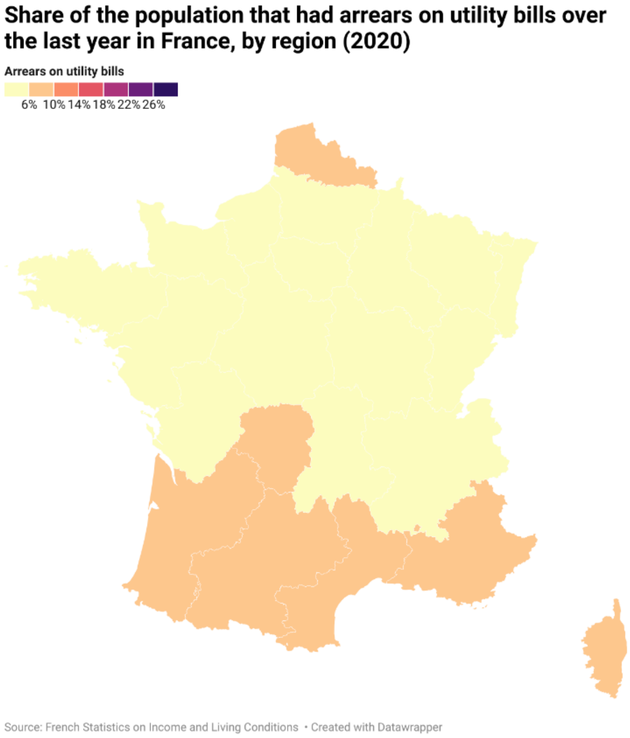

Figure 26. *Share of the French population having arrears on utility bills, by region*

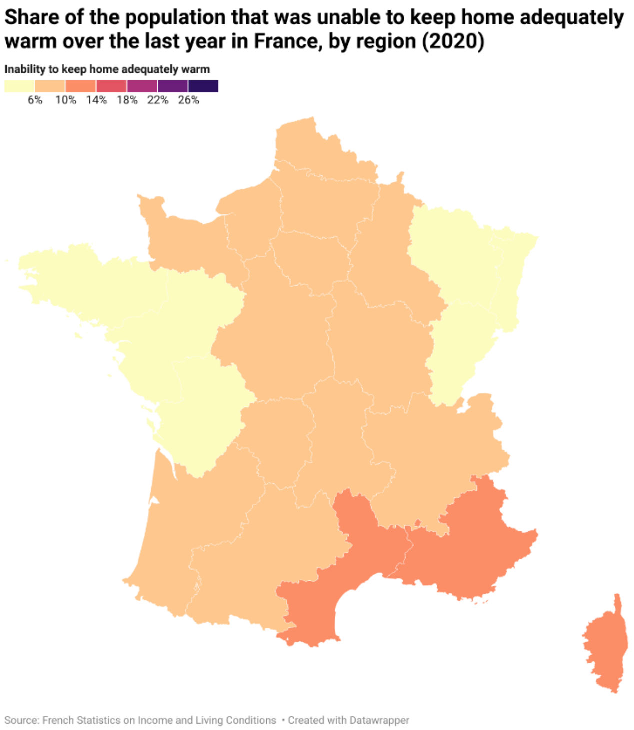

Figure 27. *Share of the French population that was unable to keep home adequately warm, by region*

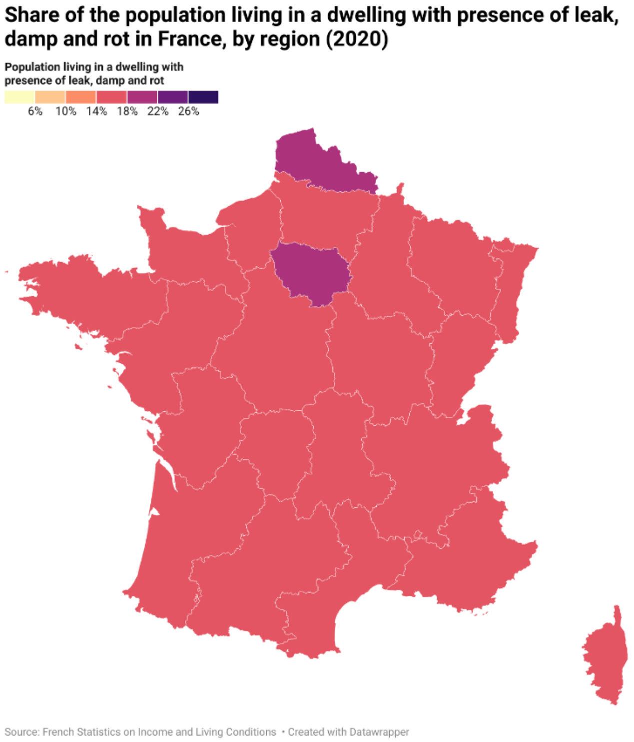

Figure 28. *Share of the French population living in a dwelling with presence of leak, damp and rot, by region*

Underconsumption, which is captured by the M/2 indicator, is more prevalent in Paris Region. Conversely, disproportionate energy expenditures, captured by the 2M indicator, is less prevalent in Paris Region.

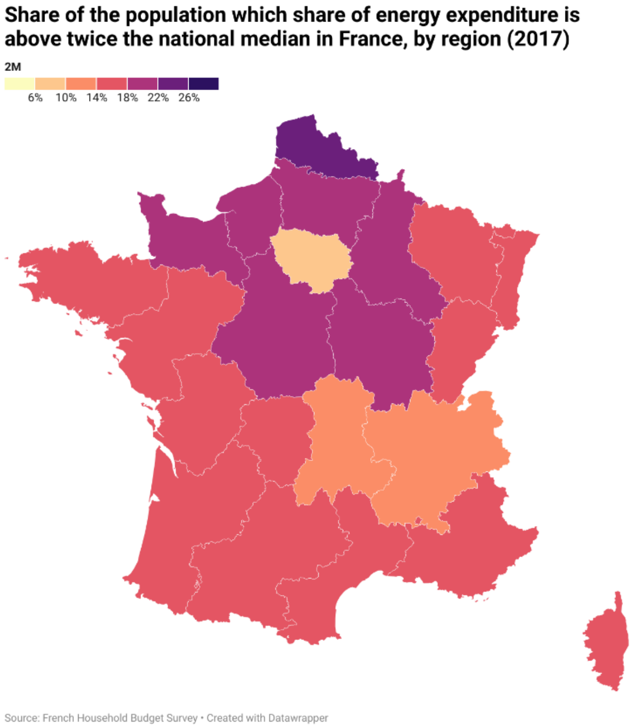

Figure 29. *Share of the French population which share of energy expenditure is above twice the national median, by region*

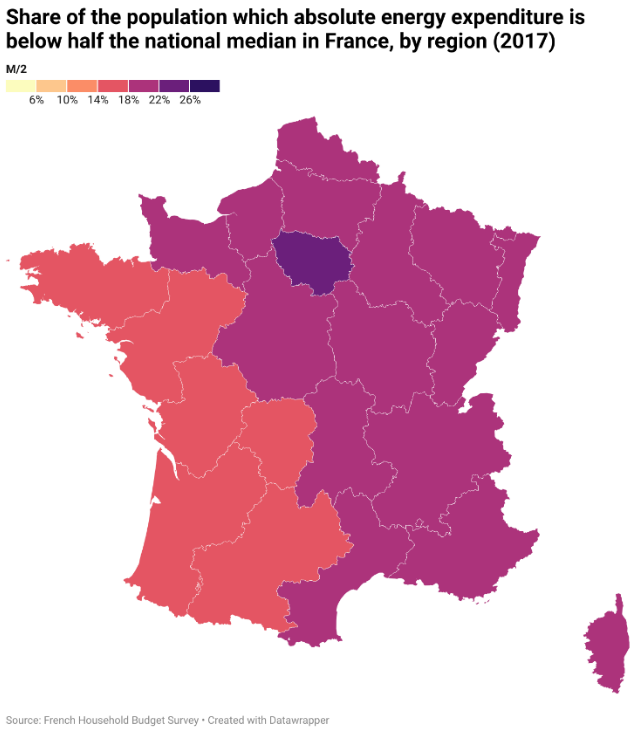

Figure 30. *Share of the French population which absolute energy expenditure is below twice the national median, by region*

IV.1.d. Comparison of energy poverty incidence in Spain, Italy and France

In this subpart, a comparative study of energy poverty incidence in Spain, Italy and France is going to be conducted for the 2017-2022 period. As we did not compute all the indicators in the three countries for this period due to lack of access to data, we completed the results obtained in the previous subsections with the indicators values from EPAH (Energy Poverty Advisory Hub, 2023). Due to the lack of annual HBS data in Italy and France and to avoid the use of proxy in the case of Italy, the comparative study was performed using only the three subjective indicators that have been computed previously. The data we are going to use are presented in the figure 32.

Spain

Indicator	2017	2018	2019	2020	2021	2022
Arrears on utility bills	9.6%	7.2%	6.6%	9.6%	9.5%	9.2%
Inability to keep home adequately warm	8.0%	9.1%	7.6%	10.9%	14.3%	17.1%
Population living in a dwelling with presence of leak, damp and rot	11.5%	15.9%	14.7%	19.7%	-	-

France

Indicator	2017	2018	2019	2020	2021	2022
Arrears on utility bills	7.5%	6.4%	5.7%	5.6%	7.1%	7.1%
Inability to keep home adequately warm	4.9%	5.0%	6.2%	6.5%	6.0%	10.9%
Population living in a dwelling with presence of leak, damp and rot	11.1%	12.6%	11.5%	17.7%	-	-

Italy

Indicator	2017	2018	2019	2020	2021	2022
Arrears on utility bills	4.1%	4.5%	4.5%	6.0%	6.5%	5.0%
Inability to keep home adequately warm	15.2%	14.1%	11.1%	8.3%	8.1%	8.8%
Population living in a dwelling with presence of leak, damp and rot	16.1%	13.2%	14.0%	19.6%	-	-

Figure 31. *Subjective energy poverty indicators in Spain, France and Italy between 2017 and 2022*

In all three countries, the situation is quite alarming, as the values of the indicators has all increased or has remained stable over the 2017-2022 period. Nevertheless, there are differences in the prevalence of the different dimensions of energy poverty in these three countries.

First, the share of the population having arrears on utility bills is higher in Spain than in France, with 9.2% of the Spanish population and 7.1% of the French population experiencing arrears in utility bills in 2022. The same indicator is lower in Italy, with a value of 5.0% in 2022. However, the three countries share similar trends. Spain and France experienced a consequent drop from 2017 and 2018, while Italian value slightly increased. Then, all three countries had a gradual increase between 2018 and 2021, followed by a small decreased in 2022.

Both Spain and France show worrying results in the inability to keep home adequately warm, an indicator that has more than doubled over the 5-year span, although the problem remains much more acute in Spain with 17.1% of the Spanish population being unable to keep home adequately warm against 10.9% of the French population in 2022. Conversely, the share of the Italian population unable to keep their home adequately warm decreased between 2019 and 2020, then it has been stable in 2021 and increased in 2022.

The main cause of concern is the surge in the share of the population living in a dwelling with presence of leaks, damp and rot in all three countries. It increased by about 50% in all three countries between 2018 and 2020 to reach 17.7% in France, 19.6% in Italy and 19.7% in Spain. This signifies a deterioration in housing quality, which could significantly impact energy efficiency, as poorly insulated homes with leaks require more energy to maintain comfortable temperatures, leading to higher energy bills.

Overall, we can also observe that the situation of energy poverty was quite "under control" between 2017 and 2019 in all three countries. However, from 2020 onwards, the situation worsened overall in these countries, especially in Spain and France. This probably reflects the impact of the COVID-19 crisis and the following energy and inflationary crisis, which have both caused a loss of purchasing power.

To sum up, despite some disparities, France, Italy and Spain share similar trends in terms of energy poverty. This reinforces the idea that harmonized policies to tackle energy poverty at the European level could be relevant, especially when looking at more affine countries.

IV.2. Identifying vulnerability factors in France

This sub-section deals with the analysis of vulnerability to energy poverty based on French SILC 2020 data. The results are presented in Figure 33. This econometric analysis aims to identify the key factors that influence the likelihood of households being energy poor in France. In this subpart, being energy poor refers to being unable to keep their home adequately warm. The model developed presents a pseudo R^2 of 0.1060.

First, the configuration of the household highly influences its likelihood to be unable to keep its home adequately warm. Couples, with or without children, are less likely to be energy poor than households formed by a single person or a single-parent family. Moreover, the larger the number of children in the household, the higher the likelihood of being energy poor. Conversely, the number of adults in the household has no observable influence on this likelihood. Thus, even though energy poverty is always linked to incomes, the composition of the households also has to be taken into account when designing measures to tackle energy poverty.

	Coefficients	Probability ratios
Dummy overpopulated household	-0.0214	0.9788
Type of household		
Single person	0.1047	1.1104
Single-parent family	-0.0464	0.9547
Couple without children	-0.6721**	0.5106**
Couple with at least one children	-0.8105**	0.4446**
Tenure status of households		
Owner	-0.4756*	0.6215*
Tenant or sub-tenant	0.2405	1.2719
Type of house		
Detached house	-0.2693	0.7639
Semi-detached house	0.0368	1.0375
Apartment in a building with less than 10 dwellings	0.1067	1.1126
Apartment in a building with 10 or more dwellings	-0.2853	0.7518
Type of employment of the main breadwinner		
Farmers	-0.1778	0.8371
Artisans, shopkeepers and company managers	0.4802	1.6164
Executives and higher intellectual professions	-0.9346**	0.3927**
Intermediate professions	0.0677	1.0700
Office employees	0.6066*	1.8342*
Workers	0.2868	1.3322
Retired people	0.1827	1.2005
Employment of the main breadwinner		
Employed	-0.8808**	0.4145**
Student or trainee	-1.3373***	0.2626***
Unemployed	-0.1406	0.8688
Retired	-0.8541***	0.4257***
Education level of the main breadwinner		
Master or equivalent	-0.9696**	0.3792**
Licence or equivalent	-0.5905	0.5541
DUT or equivalent	-0.5612	0.5705
Bac or equivalent	-0.1733	0.8409
CAP-DNB or equivalent	-0.1000	0.9048
No diploma	0.0736	1.0764
Area of residence		
High-density area	-0.1072	0.8983
Low-density area	0.2045**	1.2269**
Number of children in the household	0.1936***	1.2136***
Number of adults in the household	0.0709	1.0735

Pseudo R-squared = 0.1060
Note: Asterisks indicate the level of significance of the parameter, so that:
*** indicates significance at 1%, ** at 5%, * at 10%

Figure 32. *Vulnerability to not being able to keep home adequately warm in winter in France (2020; R^2=0.1060)*

Then, the tenure status of the household also impacts the likelihood to be unable to keep home adequately warm. The households that own their dwellings shows a much lower probability of being energy poor. Indeed, their likelihood of being unable to keep their home adequately warm is twice lower than the ones of households which rent their dwellings or are accommodated free of charge. A possible explanation for this result is that owning a dwelling is correlated with high incomes; however, this should be investigated more deeply. Furthermore, our study has not revealed any influence of the type of house on the probability of a household being energy-poor. Also, whether the dwelling is overpopulated has shown no significant influence on the likelihood of being energy poor.

It is noteworthy to note that households living in the low-density areas are more likely to be unable to keep their homes adequately warm than those living in intermediate-density or high-density areas.

Finally, when identifying energy poverty vulnerability factors for the households, the occupation of the main breadwinner has shown to be very relevant. A household is more than twice more likely to be unable to keep their home adequately warm if its main breadwinner is unemployed or professionally inactive than if he is employed, studying or retired. Additionally, working as an executive consequently reduces the probability of being energy poor. This result is easy to understand as this occupation often comes with a substantial salary. On top of that, the education level of the main breadwinner also has an influence, as having a master’s degree consequently decreases the likelihood of being energy poor. Here too, the explanation behind this result could be the higher salary to which a master's degree entitles the holder.

To sum up, non-owner households with many children and/or with one adult, living in a low-density area, with a main breadwinner in the situation of job instability, are the most vulnerable households to being unable to keep home adequately warm in France. Therefore, policies aiming to tackle this dimension of energy poverty should be primarily focused on these households. However, the low pseudo R^2 of the model obtained reminds us that energy poverty is a complex phenomenon and that it is difficult to precisely identify vulnerable households.

V. CONCLUSION

This chapter will conclude the study by summarizing the key research findings in relation to the research objectives and questions and discussing their value and contribution. It will also review the limitations of the study and suggest opportunities for future research.

V.1. Key findings

This study aimed to critically analyze the state of the art in energy poverty measurement to then develop a tool to measure energy poverty in three different European countries, namely Spain, France and Italy.

The comparative study of the incidence of energy poverty in Spain, Italy and France from 2017 to 2022 reveals a worrying situation, with all countries experiencing an increase or stability in energy poverty indicators. The prevalence of the different dimensions of energy poverty varies between these countries, with Spain showing higher disproportionate energy expenditure (2M) and inability to keep home adequately warm, while France has a higher prevalence of low absolute energy expenditure (M/2). Italy, on the other hand, has lower rates of arrears on utility bills. Despite these differences, these countries share common trends over the period of study. All three countries have seen an alarming increase in the proportion of the population living in dwellings with the presence of leaks, dampness and rot, indicating a deterioration in housing quality and potential implications for energy efficiency. Also, the energy poverty situation seemed to be improving in all three countries until 2019 but has since worsened.

To sum up, despite some discrepancies, France, Italy and Spain share similar trends in terms of energy poverty. This reinforces the idea that harmonized policies to tackle energy poverty at the European level could be relevant. This unified approach could enable better coordination, collaboration, and consistency in fighting energy poverty, as demonstrated in the past for other neighboring countries (Palma, Barrella, Gouveia, & Romero, 2024). It could involve common energy efficiency standards for housing, regulations to prevent disconnections and arrears, and measures to improve energy affordability for low-income households.

The other objective of this research was to evaluate the impact of socio-economic factors on energy poverty, with the aim of identifying the key characteristics of the households most vulnerable to energy poverty.

The econometric analysis performed has shown that non-owner households with many children and/or with one adult, living in a rural area, with a main breadwinner in the situation of job instability are more likely to be unable to keep their home adequately warm in France.

Therefore, French policies to deal with this dimension of energy poverty should target these households. In this way, these measures would be more effective and mitigate the issue.

V.2. Limitations and recommendations

The present study has provided a substantial understanding of the current circumstances of energy poverty in France, Italy and Spain, as well as the elements that contribute to vulnerability to energy poverty. However, it is crucial to recognize and discuss several limitations inherent to this research, which should be addressed in future studies to further improve the knowledge in energy poverty.

Firstly, measuring energy poverty poses a challenge in its measurement due to its multidimensional nature. While the current investigation employed five distinct indicators to evaluate energy poverty in the aforementioned countries, as explained in the Literature Review section, these indicators may not entirely and exclusively capture the entire population in situation of energy poverty. This comes from the complexity of energy poverty, which involves aspects such as household income, energy efficiency of dwellings, energy prices, and demographic factors, among others. Therefore, future research could concentrate on refining and expending the set of indicators used to measure energy poverty, enabling a more precise and comprehensive identification of affected households.

In addition, to foster a more extensive comprehension of energy poverty across Europe, it is recommended that future studies conduct comparative analyses involving other European countries. This approach would not only contribute to a more holistic understanding of energy poverty but also allow for the identification of potential regional disparities and specific factors influencing energy poverty in different European contexts.

Regarding the vulnerability analysis conducted in this study, it is important to note that it relied exclusively on 2020 French data and used the “inability to keep home adequately warm” as the sole indicator of energy poverty. While this indicator offers valuable insights into one dimension of energy poverty, it does not provide a complete picture of the various aspects of this issue. To address this limitation and to gain a more comprehensive understanding of vulnerability factors in France, it is suggested that future studies might apply the proposed vulnerability analysis’ methodology to data from multiple years and incorporate a range of energy poverty indicators as dependent variables. This approach would enable the identification of households vulnerable to the different dimensions of energy poverty in France, as well as in other countries.

Furthermore, to facilitate cross-country comparisons and enhance our understanding of vulnerability factors in diverse national contexts, it is suggested that similar studies should be conducted in other European countries. This would allow for the identification of commonalities and differences in vulnerability factors across countries, ultimately contributing to the development of more targeted and effective policies to tackle energy poverty at the European level.

VI. REFERENCES

Antepara, I., Papada, L., Gouveia, J., Katsoulakos, N., & Kaliampakos, D. (2020, July 16). Improving Energy Poverty Measurement in Southern European Regions through Equivalization of Modeled Energy Costs . *Sustainability*.

Antunes, M., Teotónico, C., Quintal, C., & Martins, R. (2023). Energy affordability across and within 26 European countries: Insights into the prevalence and depth of problems using microeconomic data. *Energy Economics, 127*.

Baker, K., Mould, R., & Restrick, S. (2018). Rethink fuel poverty as a complex problem. *Nature Energy, 3*(8), 610-612.

Barrella, R., & Romero, J. (2023). Unveiling Hidden Energy Poverty in a Time of Crisis: A Methodological Approach for National Statistics. *Living with Energy Poverty*, pp. 41-51.

Barrella, R., Linares, J. I., Romero, J. C., Arenas, E., & Centeno, E. (2021, October). Does cash money solve energy poverty? Assessing the impact of household heating allowances in Spain. *Energy Research & Social Science, 80*.

Barrella, R., Mora Rosado, S., & Romero Mora, J. (2024). *BOLETIN DE VULNERABILIDAD N°31: El impacto de la pobreza energética en la vulnerabilidad social de la poblacion atendida por CRE en el contexto de la crisis inflacionaria.* Madrid: Cruz Roja Española.

Barrella, R., Romero, J., Linares, J., Arenas, E., Asín, M., & Centeno, E. (2022, April). The dark side of energy poverty: Who is underconsuming in Spain and why? *Energy Research & Social Science, 86*.

Brabo-Catala, L., Collins, E., & Barton, B. (2024). Key Fuel Poverty Indicators and Variables: A Systemic Literature Review. *Economics of Energy, 13*(1).

Camboni, R., Corsini, A., Miniaci, R., & Valbonesi, P. (2021). Mapping fuel poverty risk at the municipal level. A small-scale analysis of Italian Energy Performance Certificate, census and survey data. *Energy Policy, 150*.

Cameron, A. C., & Trivedi, P. K. (2005). *Microeconomics: Methods and Applications.* Cambridge University Press.

Energy Poverty Advisory Hub. (2023). *National Indicators*. Retrieved June 13, 2024, from Energy Poverty Advisory Hub: https://energy-poverty.ec.europa.eu/observing-energy-poverty/national-indicators_en

EPAH. (2022). *Energy Poverty National Indicators: Insights for a more effective measuring.*

EPAH. (2023). *Energy Poverty National Indicators: Uncovering New Possibilities for Expended Knowledge.*

Eurostat. (2021). *Methodological Guidelines and Description of EU-SILC Target Variables.*

Eurostat. (2023, September 13). *9% of EU population unable to keep home warm in 2022*. Retrieved from Eurostat: https://ec.europa.eu/eurostat/web/products-eurostat-news/w/ddn-20230911-1

INE. (2019). Encuesta de presupuestos familiares - 2019.

INE. (2019). Life Conditions Survey - 2019.

INE. (2020). Encuesta de presupuestos familiares - 2020.

INE. (2020). Life Conditions Survey - 2020.

INE. (2021). Encuesta de presupuestos familiares - 2021.

INE. (2021). Life Conditions Survey - 2021.

INE. (2022). Encuesta de presupuestos familiares - 2022.

INE. (2022). Life Conditions Survey - 2022.

INSEE. (2017). Budget des familles - 2017.

INSEE. (2018). Statistiques sur les ressources et les conditions de vie (SRCV) - 2018.

INSEE. (2019). Statistiques sur les ressources et les conditions de vie (SRCV) - 2019.

INSEE. (2020). Statistiques sur les ressources et les conditions de vie (SRCV) - 2020.

International Energy Agency. (2023). *Italy 2023 Energy Policy Review.*

ISTAT. (2019). Household Budget Survey: public use microdata - 2019.

ISTAT. (2020). Household Budget Survey: public use microdata - 2020.

ISTAT. (2021). Household Budget Survey: public use microdata - 2021.

ISTAT. (2022). Household Budget Survey: public use microdata - 2022.

ISTAT. (2023, September 29). *Microdata Files*. Retrieved from ISTAT: https://www.istat.it/en/analysis-and-products/microdata-files

Koukoufikis, G., & Uihlein, A. (2022). *Energy poverty, transport poverty and living conditions-An analysis of EU data and socioeconomic indicators.*

Legendre, B., & Ricci, O. (2015, May). Measuring fuel poverty in France: Which households are the most fuel vulnerable? *Energy Economics, 49*, pp. 620-628.

Menyhert, B. (2023). Existing chalenges and potential improvements to energy poverty measurement. *EPAH Annual Conference.* Warsow.

Meyer, S., Laurence, H., Bart, D., Middlemiss, L., & Maréchal, K. (2018, June). Capturing the multifaceted nature of energy poverty: Lessons from Belgium. *Energy Research & Social Science, 40*, pp. 273-283.

Palma, P., Barrella, R., Gouveia, J. P., & Romero, J. C. (2024, October). Comparative analysis of energy poverty definition and measurement in Portugal and Spain. *Utilities Policy, 90*.

Romero, J., Centeno Hernáez, E., & Barrella, R. (2022). *Informe de Indicadores de Pobreza Energética en España 2021.*

Romero, J., Linares, P., & López, X. (2018, April). The policy implications of energy poverty indicators. *Energy Policy, 115*, pp. 98-108.

Romero, J., Linares, P., & López-Otero, X. (2017, July). The policy implications of energy poverty indicators.

Thema, J., & Vondung, F. (2020). EPOV indicator dashboard: Methodology guidebook. *Wuppertal Institut für Klima, Umwelt, Energie GmbH.*

VII. APPENDICES

Appendix 1: Alignment with Sustainable Development Goals

Three different Sustainable Development Goals outlined by the United Nations are particularly relevant to this project:

- **Goal 1: No Poverty**. By developing a framework and monitoring tool to characterize energy poverty in Europe, the research contributes to understanding and mitigating a certain type of poverty, which is energy poverty.
- **Goal 3: Good Health and Well-being**. Having access to sufficient energy is necessary to good health and well-being. Then, addressing energy poverty issues aims to ensure healthy lives and promote well-being.
- **Goal 7: Affordable and Clean Energy**. By evaluating the impact of socio-economic factors on energy poverty and designing a measurement tool, the project might lead to the development of policies to favor access to affordable and reliable energy.

Appendix 2: Vulnerability Analysis Jupyter Notebook

Analysis of vulnerability to energy poverty

Clément Baumann

Data preparation

```
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np
import statsmodels.api as sm
import seaborn as sns
import scipy.stats as stats
```

```
# Import data from the French 2020 SILC
df = pd.read_csv('Data/France/SRCV2020/Csv/menages20_diff.csv', sep=';')
df
```

```
C:\Users\clemb\AppData\Local\Temp\ipykernel_22676\1506642150.py:2: DtypeWarning: Columns (19,93,114,115,121,122,129,150,161,168,171,184,195,197,201,203,205,214,222,231,232,237,242,243,246,255,263,267,272,286,287,288,289,290,292,296,298,302,309,312,317,331,338,359,361,363,367,369,466) have mixed types. Specify dtype option on import or set low_memory=False.
  df = pd.read_csv('Data/France/SRCV2020/Csv/menages20_diff.csv', sep=';')
```

```
       IDMEN    DB030  BAIN  BAIN_DRAP  BRUIT  BRUIT_DRAP   CALIM  CALIM_DRAP  \
0          1  4351300   1.0          1    2.0           1   600.0           1
1          2  4351400   1.0          1    2.0           1   960.0           1
2          3  4351500   1.0          1    2.0           1   400.0           1
3          4  4352500   1.0          1    2.0           1  1500.0           1
4          5  4352700   1.0          1    2.0           1   520.0           1
...      ...      ...   ...        ...    ...         ...     ...         ...
10894  10895  6442500   1.0          1    2.0           1   900.0           1
10895  10896  6442600   1.0          1    2.0           1   600.0           1
10896  10897  6442900   1.0          1    2.0           1   800.0           1
10897  10898  6443100   1.0          1    2.0           1   600.0           1
10898  10899  6443400   1.0          1    2.0           1   150.0           1

       CRESTO  CRESTO_DRAP  ...       IDPREF      IDRQM       IDPRRP MODECOLL  \
0       250.0            1  ...  435130001.0  435130001  435130001.0      FAF
1       300.0            1  ...  435140001.0  435140001  435140001.0      FAF
2       150.0            1  ...  435150001.0  435150002  435150002.0      TEL
3      1000.0            1  ...  435250001.0  435250001  435250001.0      TEL
4       420.0            1  ...  435270001.0  435270001  435270002.0      FAF
...       ...          ...  ...          ...        ...          ...      ...
10894    50.0            1  ...  644250002.0  644250001  644250002.0      TEL
10895     0.0            1  ...  644260001.0  644260001  644260001.0      TEL
10896   200.0            1  ...  644290002.0  644290001  644290002.0      FAF
10897   150.0            1  ...  644310002.0  644310001  644310001.0      FAF
10898     0.0            1  ...  644340001.0  644340001  644340001.0      TEL
```

```
       HX050  HX080          HX090  WORK_INT  WORK_INT_NEW PLG_QP
0        1.0  False   28382.425383       NaN           NaN      0
1        1.5  False   36829.507182       NaN           NaN      0
2        1.5   True -15379.976500  1.000000      1.000000      0
3        2.5  False   71223.447840  1.000000      1.000000      0
4        2.5  False   35410.845362  1.000000      1.000000      0
...      ...    ...            ...       ...           ...    ...
10894    2.0  False   23053.211313  0.166667      0.083333      0
10895    1.8  False   23580.508793  0.625000      0.625000      0
10896    3.0  False   24665.913860  1.000000      1.000000      0
10897    2.5  False   28859.823970  1.000000      1.000000      0
10898    1.0   True   11485.978462       NaN           NaN      0

[10899 rows x 467 columns]

# select the columns under study
data=pd.DataFrame()

data['ArrearsOnUtilityBills'] = df['IPELEC'].copy()
data['UnableToKeepHomeWarm'] = df['TEMP'].copy()
data['DwellingWithLeakingRoof'] = df['TOIT'].copy()

data['HouseholdType'] = df['TYPMEN5'].copy()
data['TenureStatus'] = df['STO'].copy()
data['HouseType'] = df['TYPLOG'].copy()
data['EmploymentType'] = df['CS24PR'].copy()
data['Employment'] = df['SITUAPR'].copy()
data['EducationLevel'] = df['DIPDETPR'].copy()
data['AreaOfResidence'] = df['DB100'].copy()
data['NbChildren'] = df['NENFANTS'].copy()
data['NbAdults'] = df['NPERS'].copy() - df['NENFANTS'].copy()
data['NbAdults'][data['NbAdults'] < 0] = 0
data['Overpopulated'] = df['DENS'].copy()

data

C:\Users\clemb\AppData\Local\Temp\ipykernel_22676\1177811915.py:17: 
SettingWithCopyWarning: 
A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation: https://pandas.pydata.org/pandas-
docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  data['NbAdults'][data['NbAdults'] < 0] = 0

       ArrearsOnUtilityBills  UnableToKeepHomeWarm  DwellingWithLeakingRoof  \
0                        3.0                   1.0                      2.0   
1                        3.0                   1.0                      2.0   
2                        3.0                   1.0                      2.0   
3                        3.0                   1.0                      2.0   
4                        3.0                   1.0                      2.0   
...                      ...                   ...                      ...   
10894                    3.0                   1.0                      1.0   
10895                    3.0                   1.0                      2.0   
10896                    3.0                   1.0                      2.0   
```

```
10897                    3.0                   1.0                        2.0
10898                    3.0                   1.0                        2.0

       HouseholdType  TenureStatus  HouseType  EmploymentType  Employment  \
0                  1             2          3            73.0         5.0
1                  3             1          3            73.0         5.0
2                  3             2          4            32.0         1.0
3                  4             1          4            36.0         1.0
4                  4             1          4            36.0         1.0
...              ...           ...        ...             ...         ...
10894              5             1          2            73.0         5.0
10895              4             1          2            55.0         1.0
10896              4             1          2            36.0         1.0
10897              4             1          1            36.0         1.0
10898              1             1          2            76.0         5.0

       EducationLevel AreaOfResidence  NbChildren  NbAdults  Overpopulated
0                22.0               1           0         1              0
1                11.0               1           0         2              0
2                17.0               1           0         2              0
3                13.0               1           2         2              0
4                32.0               1           2         2              0
...               ...             ...         ...       ...            ...
10894            51.0               2           0         3              0
10895            52.0               2           1         2              0
10896            32.0               2           3         2              1
10897            16.0               2           2         2              0
10898            71.0               3           0         1              0

[10899 rows x 13 columns]

data.describe()

       ArrearsOnUtilityBills  UnableToKeepHomeWarm  DwellingWithLeakingRoof  \
count           10876.000000          10865.000000             10762.000000
mean                2.947315              1.070502                 1.831723
std                 0.330630              0.256002                 0.374130
min                 1.000000              1.000000                 1.000000
25%                 3.000000              1.000000                 2.000000
50%                 3.000000              1.000000                 2.000000
75%                 3.000000              1.000000                 2.000000
max                 4.000000              2.000000                 2.000000

       HouseholdType  TenureStatus     HouseType  EmploymentType  \
count   10899.000000  10899.000000  10899.000000    10853.000000
mean        2.572805      1.363336      2.147903       57.380816
std         1.254355      0.521090      1.271090       17.450655
min         1.000000      1.000000      1.000000       10.000000
25%         1.000000      1.000000      1.000000       41.000000
50%         3.000000      1.000000      2.000000       61.000000
75%         4.000000      2.000000      3.000000       73.000000
max         5.000000      3.000000      9.000000       82.000000

         Employment  EducationLevel    NbChildren      NbAdults  Overpopulated
count  10898.000000    10885.000000  10899.000000  10899.000000   10899.000000
```

```
mean       2.779409      45.063666      0.675750      1.595834      0.062942
std        2.039882      17.997138      1.032599      0.522098      0.242869
min        1.000000      11.000000      0.000000      0.000000      0.000000
25%        1.000000      32.000000      0.000000      1.000000      0.000000
50%        1.000000      51.000000      0.000000      2.000000      0.000000
75%        5.000000      53.000000      1.000000      2.000000      0.000000
max        8.000000      71.000000      8.000000      5.000000      1.000000
```

```
# number of NaN values in the dataframe
data.isnull().sum()
```

```
ArrearsOnUtilityBills       23
UnableToKeepHomeWarm        34
DwellingWithLeakingRoof    137
HouseholdType                0
TenureStatus                 0
HouseType                    0
EmploymentType              46
Employment                   1
EducationLevel              14
AreaOfResidence              0
NbChildren                   0
NbAdults                     0
Overpopulated                0
dtype: int64
```

```
# delete rows with NaN values
data.dropna(axis=0, inplace=True)
data.reset_index(drop=True, inplace=True)
data
```

```
       ArrearsOnUtilityBills  UnableToKeepHomeWarm  DwellingWithLeakingRoof  \
0                        3.0                   1.0                      2.0
1                        3.0                   1.0                      2.0
2                        3.0                   1.0                      2.0
3                        3.0                   1.0                      2.0
4                        3.0                   1.0                      2.0
...                      ...                   ...                      ...
10668                    3.0                   1.0                      1.0
10669                    3.0                   1.0                      2.0
10670                    3.0                   1.0                      2.0
10671                    3.0                   1.0                      2.0
10672                    3.0                   1.0                      2.0

       HouseholdType  TenureStatus  HouseType  EmploymentType  Employment  \
0                  1             2          3            73.0         5.0
1                  3             1          3            73.0         5.0
2                  3             2          4            32.0         1.0
3                  4             1          4            36.0         1.0
4                  4             1          4            36.0         1.0
...              ...           ...        ...             ...         ...
10668              5             1          2            73.0         5.0
10669              4             1          2            55.0         1.0
10670              4             1          2            36.0         1.0
10671              4             1          1            36.0         1.0
10672              1             1          2            76.0         5.0
```

```
       EducationLevel AreaOfResidence  NbChildren  NbAdults  Overpopulated
0                22.0               1           0         1              0
1                11.0               1           0         2              0
2                17.0               1           0         2              0
3                13.0               1           2         2              0
4                32.0               1           2         2              0
...               ...             ...         ...       ...            ...
10668            51.0               2           0         3              0
10669            52.0               2           1         2              0
10670            32.0               2           3         2              1
10671            16.0               2           2         2              0
10672            71.0               3           0         1              0

[10673 rows x 13 columns]
```

```
# for the next three variable, I transform them into boolean variables
data['ArrearsOnUtilityBills'] = data['ArrearsOnUtilityBills'] <= 2     # TRUE if 
there are arrears
data['UnableToKeepHomeWarm'] = data['UnableToKeepHomeWarm'] == 2       # TRUE if 
the houshold is unable to keep home warm
data['DwellingWithLeakingRoof'] = data['DwellingWithLeakingRoof'] == 1 # TRUE if 
there is a leaking roof

data['Overpopulated'] = data['Overpopulated'] == 1   #Transformation in a boolean 
variable; TRUE if the dwelling is overpopulated

# For the next variables, I use dictionnaries of values
# in the microdata, the values are number ; the correspondant values are given in 
explanotory documents

dic = {1: 'Personne seule', 2: 'Famille monoparentale', 3: 'Couple sans enfant', 
4: 'Couple avec au moins un enfant', 5: 'Autre'}
data['HouseholdType'] = data['HouseholdType'].transform(lambda x: dic[x])

dic = {1: 'Propriétaire', 2: 'Locataire ou sous-locataire', 3: 'Logé à titre 
gratuit'}
data['TenureStatus'] = data['TenureStatus'].transform(lambda x: dic[x])

#dic = {1: 'Maison indépendante', 2: 'Maison mitoyenne', 3: 'Appartement dans un 
immeuble de moins de 10 logements',
#       4: 'Appartement dans un immeuble 10 logements ou plus', 5: 'Logement dans 
un foyer pour personnes âgées, étudiants, travailleurs…',
#       6: 'Logement dans un immeuble collectif à usage autre que d’habitation', 
7: 'Chambre d’hôtel',
#       8: 'Construction provisoire', 9: 'Autre type de logement'}
dic = {1: 'Maison indépendante', 2: 'Maison mitoyenne', 3: 'Appartement dans un 
immeuble de moins de 10 logements',
       4: 'Appartement dans un immeuble 10 logements ou plus', 5: 'Autre type de 
logement',
       6: 'Autre type de logement', 7: 'Autre type de logement',
       8: 'Autre type de logement', 9: 'Autre type de logement'}    # Here, I have 
reduced the number of values possible
data['HouseType'] = data['HouseType'].transform(lambda x: dic[x])
```

```
dic = {10: 'Agriculteurs exploitants',
       21: 'Artisans, commerçants et chefs d’entreprise', 22: 'Artisans, commerçants et chefs d’entreprise', 23: "Artisans, commerçants et chefs d’entreprise",
       31: 'Cadres et professions intellectuelles supérieures', 32: 'Cadres et professions intellectuelles supérieures', 36: "Cadres et professions intellectuelles supérieures",
       41: "Professions intermédiaires", 46: "Professions intermédiaires", 47: "Professions intermédiaires", 48: "Professions intermédiaires",
       51: "Employés", 54: "Employés", 55: "Employés", 56: "Employés",
       61: "Ouvriers", 66: "Ouvriers", 69: "Ouvriers",
       71: "Retraités", 72: "Retraités", 73: "Retraités", 76: "Retraités",
       81: "Autres personnes sans activité professionnelle", 82: "Autres personnes sans activité professionnelle"}
data['EmploymentType'] = data['EmploymentType'].transform(lambda x: dic[x])

#dic = {1:" Occupe un emploi", 2: "Apprenti(e) sous contrat ou stagiaire rémunéré(e)", 3: "Etudiant(e), élève, en formation ou stagiaire non rémunéré(e)",
#       4: "Au chômage (inscrit(e) ou non au Pôle Emploi (ex ANPE))", 5: "Retraité(e) ou retiré(e) des affaires ou en préretraite",
#       6: "Femme ou homme au foyer", 7: "Inactif(ve) pour cause d'invalidité", 8: "Autre situation d'inactivité"}
dic = {1:"Occupe un emploi", 2: "Etudiant, en formation ou stagiaire", 3: "Etudiant, en formation ou stagiaire",
       4: "Au chômage", 5: "Retraité(e)",
       6: "Inactif(ve)", 7: "Inactif(ve)", 8: "Inactif(ve)"}       # Here, I have reduced the number of values possible
data['Employment'] = data['Employment'].transform(lambda x: dic[x])

#dic = {11: "Doctorats sauf santé", 12: "Doctorats de santé", 13: "Ecoles de commerce", 14: "Ecoles d'ingénieur", 15: "DESS, masters professionnels",
#       16: "DEA, Magistères, master recherche", 17: "Master non différencié", 21: "Maîtrise, MST, Miage, maîtrise IUP", 22: "Licence",
#       23: "Licence professionnelle, licence IUP", 24: "Autres diplômes supérieurs (niveau bac +3 et plus)", 31: "DUT", 32: "BTS",
#       33: "Deust, DTS, DNTS, DPECF", 34: "Diplômes paramédicaux et sociaux (niveau bac +2)", 35: "Deug", 36: "Autres diplômes niveau technicien supérieur (niveau bac +2)",
#       41: "Capacité en droit, DAEU, ESEU", 42: "Bac général", 43: "Bac technologique", 44: "Baccalauréat professionnel",
#       45: "Brevet de technicien, brevet professionnel", 51: "BEP", 52: "CAP", 53: "Autres diplômes de niveau CAP-BEP",
#       60: "Brevet des collèges", 70: "Certificat d'études primaires", 71: "Aucun diplôme"}
dic = {11: "Doctorat ou équivalent", 12: "Doctorat ou équivalent", 13: "Master ou équivalent", 14:  "Master ou équivalent", 15:  "Master ou équivalent",
       16:  "Master ou équivalent", 17:  "Master ou équivalent", 21: "Licence ou équivalent", 22: "Licence ou équivalent",
       23: "Licence ou équivalent", 24: "Licence ou équivalent", 31: "DUT ou équivalent", 32: "DUT ou équivalent",
       33: "DUT ou équivalent", 34: "DUT ou équivalent", 35: "DUT ou équivalent", 36: "DUT ou équivalent",
       41: "Bac ou équivalent", 42: "Bac ou équivalent", 43: "Bac ou équivalent", 44: "Bac ou équivalent",
       45: "Bac ou équivalent", 51: "CAP-DNB ou équivalent", 52: "CAP-DNB ou
```

```
équivalent", 53: "CAP-DNB ou équivalent",
       60: "CAP-DNB ou équivalent", 70: "Sans diplôme", 71:"Sans diplôme"}
data['EducationLevel'] = data['EducationLevel'].transform(lambda x: dic[x])    #
Here, I have reduced the number of values possible

data['AreaOfResidence'] = data['AreaOfResidence'].astype('int64')
dic = {1: "Zone à forte densité de population", 2: "Zone à densité intermédiaire",
3: "Zone à faible densité de population"}
data['AreaOfResidence'] = data['AreaOfResidence'].transform(lambda x: dic[x])

data

       ArrearsOnUtilityBills  UnableToKeepHomeWarm  DwellingWithLeakingRoof  \
0                      False                 False                    False
1                      False                 False                    False
2                      False                 False                    False
3                      False                 False                    False
4                      False                 False                    False
...                      ...                   ...                      ...
10668                  False                 False                     True
10669                  False                 False                    False
10670                  False                 False                    False
10671                  False                 False                    False
10672                  False                 False                    False

                    HouseholdType                 TenureStatus  \
0                  Personne seule  Locataire ou sous-locataire
1               Couple sans enfant                 Propriétaire
2               Couple sans enfant  Locataire ou sous-locataire
3       Couple avec au moins un enfant                 Propriétaire
4       Couple avec au moins un enfant                 Propriétaire
...                           ...                          ...
10668                       Autre                 Propriétaire
10669   Couple avec au moins un enfant                 Propriétaire
10670   Couple avec au moins un enfant                 Propriétaire
10671   Couple avec au moins un enfant                 Propriétaire
10672                  Personne seule                 Propriétaire

                                              HouseType  \
0       Appartement dans un immeuble de moins de 10 lo...
1       Appartement dans un immeuble de moins de 10 lo...
2       Appartement dans un immeuble 10 logements ou plus
3       Appartement dans un immeuble 10 logements ou plus
4       Appartement dans un immeuble 10 logements ou plus
...                                                 ...
10668                                   Maison mitoyenne
10669                                   Maison mitoyenne
10670                                   Maison mitoyenne
10671                                Maison indépendante
10672                                   Maison mitoyenne

                                     EmploymentType        Employment  \
0                                         Retraités      Retraité(e)
1                                         Retraités      Retraité(e)
2       Cadres et professions intellectuelles supérieures  Occupe un emploi
```

```
3      Cadres et professions intellectuelles supérieures  Occupe un emploi
4      Cadres et professions intellectuelles supérieures  Occupe un emploi
...                                                ...               ...
10668                                           Retraités      Retraité(e)
10669                                            Employés  Occupe un emploi
10670  Cadres et professions intellectuelles supérieures  Occupe un emploi
10671  Cadres et professions intellectuelles supérieures  Occupe un emploi
10672                                           Retraités      Retraité(e)

               EducationLevel                      AreaOfResidence  \
0       Licence ou équivalent   Zone à forte densité de population
1      Doctorat ou équivalent   Zone à forte densité de population
2        Master ou équivalent   Zone à forte densité de population
3        Master ou équivalent   Zone à forte densité de population
4           DUT ou équivalent   Zone à forte densité de population
...                       ...                                  ...
10668   CAP-DNB ou équivalent         Zone à densité intermédiaire
10669   CAP-DNB ou équivalent         Zone à densité intermédiaire
10670       DUT ou équivalent         Zone à densité intermédiaire
10671    Master ou équivalent         Zone à densité intermédiaire
10672            Sans diplôme  Zone à faible densité de population

       NbChildren  NbAdults  Overpopulated
0               0         1          False
1               0         2          False
2               0         2          False
3               2         2          False
4               2         2          False
...           ...       ...            ...
10668           0         3          False
10669           1         2          False
10670           3         2           True
10671           2         2          False
10672           0         1          False

[10673 rows x 13 columns]
```

```
data['ArrearsOnUtilityBills'] = data['ArrearsOnUtilityBills'].astype('int')
data['UnableToKeepHomeWarm'] = data['UnableToKeepHomeWarm'].astype('int')
data['DwellingWithLeakingRoof'] = data['DwellingWithLeakingRoof'].astype('int')
data['HouseholdType'] = data['HouseholdType'].astype('category')
data['TenureStatus'] = data['TenureStatus'].astype('category')
data['HouseType'] = data['HouseType'].astype('category')
data['EmploymentType'] = data['EmploymentType'].astype('category')
data['Employment'] = data['Employment'].astype('category')
data['EducationLevel'] = data['EducationLevel'].astype('category')
data['AreaOfResidence'] = data['AreaOfResidence'].astype('category')
data['Overpopulated'] = data['Overpopulated'].astype('int')
data['NbChildren'] = data['NbChildren'].astype('float64')
data['NbAdults'] = data['NbAdults'].astype('float64')
```

```
data.dtypes
```

```
ArrearsOnUtilityBills         int32
UnableToKeepHomeWarm          int32
DwellingWithLeakingRoof       int32
```

```
HouseholdType          category
TenureStatus           category
HouseType              category
EmploymentType         category
Employment             category
EducationLevel         category
AreaOfResidence        category
NbChildren              float64
NbAdults                float64
Overpopulated             int32
dtype: object
```

Analysis of the vulnerability of households to being unable to keep home adequately warm

```
# Transforming categorical variable into dummies
df_hsld_type = pd.get_dummies(data["HouseholdType"], dtype=int)
df_hsld_type.drop(columns=['Autre'], inplace=True)

df_hsld_type.head()

   Couple avec au moins un enfant  Couple sans enfant  Famille monoparentale  \
0                               0                   0                      0
1                               0                   1                      0
2                               0                   1                      0
3                               1                   0                      0
4                               1                   0                      0

   Personne seule
0               1
1               0
2               0
3               0
4               0

df_hsld_type.describe()

       Couple avec au moins un enfant  Couple sans enfant  \
count                    10673.000000        10673.000000
mean                         0.274899            0.286892
std                          0.446484            0.452332
min                          0.000000            0.000000
25%                          0.000000            0.000000
50%                          0.000000            0.000000
75%                          1.000000            1.000000
max                          1.000000            1.000000

       Famille monoparentale  Personne seule
count           10673.000000    10673.000000
mean                0.089572        0.327930
std                 0.285581        0.469481
min                 0.000000        0.000000
25%                 0.000000        0.000000
```

```
50%                  0.000000        0.000000
75%                  0.000000        1.000000
max                  1.000000        1.000000

# Transforming categorical variable into dummies
df_tenure_type = pd.get_dummies(data["TenureStatus"], dtype=int)
df_tenure_type.drop(columns=['Logé à titre gratuit'], inplace=True)


df_tenure_type.head()

   Locataire ou sous-locataire  Propriétaire
0                            1             0
1                            0             1
2                            1             0
3                            0             1
4                            0             1

# Transforming categorical variable into dummies
df_hs_type = pd.get_dummies(data["HouseType"], dtype=int)
df_hs_type.drop(columns=['Autre type de logement'], inplace=True)

df_hs_type.head()

   Appartement dans un immeuble 10 logements ou plus  \
0                                                  0
1                                                  0
2                                                  1
3                                                  1
4                                                  1

   Appartement dans un immeuble de moins de 10 logements  Maison indépendante  \
0                                                  1                    0
1                                                  1                    0
2                                                  0                    0
3                                                  0                    0
4                                                  0                    0

   Maison mitoyenne
0                 0
1                 0
2                 0
3                 0
4                 0

df_hs_type.describe()

       Appartement dans un immeuble 10 logements ou plus  \
count                                       10673.000000
mean                                            0.238733
std                                             0.426329
min                                             0.000000
25%                                             0.000000
50%                                             0.000000
75%                                             0.000000
max                                             1.000000
```

```
       Appartement dans un immeuble de moins de 10 logements  \
count                                       10673.000000
mean                                            0.104094
std                                             0.305397
min                                             0.000000
25%                                             0.000000
50%                                             0.000000
75%                                             0.000000
max                                             1.000000

       Maison indépendante  Maison mitoyenne
count         10673.000000      10673.000000
mean              0.461913          0.187764
std               0.498571          0.390542
min               0.000000          0.000000
25%               0.000000          0.000000
50%               0.000000          0.000000
75%               1.000000          0.000000
max               1.000000          1.000000
```

```
# Transforming categorical variable into dummies
df_emp_type = pd.get_dummies(data["EmploymentType"], dtype=int)
df_emp_type.drop(columns=["Autres personnes sans activité professionnelle"],
inplace=True)

df_emp_type.head()
```

```
   Agriculteurs exploitants  Artisans, commerçants et chefs d'entreprise  \
0                         0                                            0
1                         0                                            0
2                         0                                            0
3                         0                                            0
4                         0                                            0

   Cadres et professions intellectuelles supérieures  Employés  Ouvriers  \
0                                                  0         0         0
1                                                  0         0         0
2                                                  1         0         0
3                                                  1         0         0
4                                                  1         0         0

   Professions intermédiaires  Retraités
0                           0          1
1                           0          1
2                           0          0
3                           0          0
4                           0          0
```

```
df_emp_type.describe()
```

```
       Agriculteurs exploitants  Artisans, commerçants et chefs d'entreprise  \
count              10673.000000                                 10673.000000
mean                   0.010400                                     0.035604
std                    0.101454                                     0.185309
min                    0.000000                                     0.000000
```

```
25%                    0.000000                                     0.000000
50%                    0.000000                                     0.000000
75%                    0.000000                                     0.000000
max                    1.000000                                     1.000000

       Cadres et professions intellectuelles supérieures      Employés   \
count                                       10673.000000  10673.000000
mean                                            0.139792      0.128642
std                                             0.346787      0.334819
min                                             0.000000      0.000000
25%                                             0.000000      0.000000
50%                                             0.000000      0.000000
75%                                             0.000000      0.000000
max                                             1.000000      1.000000

           Ouvriers  Professions intermédiaires     Retraités
count  10673.000000                10673.000000  10673.000000
mean       0.137169                    0.157313      0.352291
std        0.344041                    0.364113      0.477706
min        0.000000                    0.000000      0.000000
25%        0.000000                    0.000000      0.000000
50%        0.000000                    0.000000      0.000000
75%        0.000000                    0.000000      1.000000
max        1.000000                    1.000000      1.000000

# Transforming categorical variable into dummies
df_emp = pd.get_dummies(data["Employment"], dtype=int)
df_emp.drop(columns=["Inactif(ve)"], inplace=True)

df_emp.head()

   Au chômage  Etudiant, en formation ou stagiaire  Occupe un emploi  \
0           0                                    0                 0
1           0                                    0                 0
2           0                                    0                 1
3           0                                    0                 1
4           0                                    0                 1

   Retraité(e)
0            1
1            1
2            0
3            0
4            0

# Transforming categorical variable into dummies
df_edu = pd.get_dummies(data["EducationLevel"], dtype=int)
df_edu.drop(columns=["Doctorat ou équivalent"], inplace=True)

df_edu.head()

   Bac ou équivalent  CAP-DNB ou équivalent  DUT ou équivalent  \
0                  0                      0                  0
1                  0                      0                  0
2                  0                      0                  0
3                  0                      0                  0
```

```
4                   0                      0                  1

   Licence ou équivalent  Master ou équivalent  Sans diplôme
0                      1                     0             0
1                      0                     0             0
2                      0                     1             0
3                      0                     1             0
4                      0                     0             0

df_edu.describe()

       Bac ou équivalent  CAP-DNB ou équivalent  DUT ou équivalent  \
count       10673.000000           10673.000000       10673.000000
mean            0.157125               0.309566           0.132671
std             0.363936               0.462336           0.339235
min             0.000000               0.000000           0.000000
25%             0.000000               0.000000           0.000000
50%             0.000000               0.000000           0.000000
75%             0.000000               1.000000           0.000000
max             1.000000               1.000000           1.000000

       Licence ou équivalent  Master ou équivalent  Sans diplôme
count           10673.000000          10673.000000  10673.000000
mean                0.105968              0.076455      0.193573
std                 0.307812              0.265737      0.395116
min                 0.000000              0.000000      0.000000
25%                 0.000000              0.000000      0.000000
50%                 0.000000              0.000000      0.000000
75%                 0.000000              0.000000      0.000000
max                 1.000000              1.000000      1.000000

# Transforming categorical variable into dummies
df_res = pd.get_dummies(data["AreaOfResidence"], dtype=int)
df_res.drop(columns=["Zone à densité intermédiaire"], inplace=True)

df_res.head()

   Zone à faible densité de population  Zone à forte densité de population
0                                    0                                   1
1                                    0                                   1
2                                    0                                   1
3                                    0                                   1
4                                    0                                   1

#concatenate all the explanatory variables
df_global = data[['Overpopulated','NbAdults','NbChildren']]
x = pd.concat([df_global, df_hsld_type, df_tenure_type, df_hs_type, df_emp_type,
df_emp, df_edu, df_res], axis = 1)

y = data['UnableToKeepHomeWarm']
x = sm.add_constant(x) #  we need to add the intercept term explicitly

#fit linear regression model
model = sm.Logit(y, x)
regr_results_model = model.fit()
```

```
print(regr_results_model.summary())

Optimization terminated successfully.
         Current function value: 0.227170
         Iterations 8
                            Logit Regression Results
================================================================================
Dep. Variable:     UnableToKeepHomeWarm   No. Observations:                10673
Model:                            Logit   Df Residuals:                    10640
Method:                             MLE   Df Model:                           32
Date:                  Tue, 18 Jun 2024   Pseudo R-squ.:                  0.1060
Time:                          10:30:15   Log-Likelihood:                -2424.6
converged:                         True   LL-Null:                       -2712.0
Covariance Type:              nonrobust   LLR p-value:                9.359e-101
==================================================================================
=======================================
                                                                  coef    std err
z       P>|z|      [0.025      0.975]
----------------------------------------------------------------------------------
---------------------------------------
const                                                          -1.4624      1.051      -
1.391       0.164      -3.522       0.598
Overpopulated                                                  -0.0214      0.145      -
0.148       0.882      -0.305       0.262
NbAdults                                                        0.0709      0.313
0.227       0.821      -0.542       0.683
NbChildren                                                      0.1936      0.062
3.143       0.002       0.073       0.314
Couple avec au moins un enfant                                 -0.8105      0.331      -
2.449       0.014      -1.459      -0.162
Couple sans enfant                                             -0.6721      0.325      -
2.065       0.039      -1.310      -0.034
Famille monoparentale                                          -0.0464      0.577      -
0.080       0.936      -1.177       1.084
Personne seule                                                  0.1047      0.570
0.184       0.854      -1.012       1.222
Locataire ou sous-locataire                                     0.2405      0.254
0.946       0.344      -0.258       0.739
Propriétaire                                                   -0.4756      0.257      -
1.853       0.064      -0.979       0.027
Appartement dans un immeuble 10 logements ou plus              -0.2853      0.378      -
0.754       0.451      -1.027       0.456
Appartement dans un immeuble de moins de 10 logements           0.1067      0.383
0.278       0.781      -0.645       0.858
Maison indépendante                                            -0.2693      0.390      -
0.691       0.489      -1.033       0.494
Maison mitoyenne                                                0.0368      0.385
0.096       0.924      -0.718       0.791
Agriculteurs exploitants                                       -0.1778      0.617      -
0.288       0.773      -1.388       1.032
Artisans, commerçants et chefs d'entreprise                     0.4802      0.390
1.230       0.219      -0.285       1.246
Cadres et professions intellectuelles supérieures              -0.9346      0.426      -
2.194       0.028      -1.770      -0.100
Employés                                                        0.6066      0.347
```

```
1.747      0.081      -0.074      1.287
Ouvriers                                                     0.2868      0.354
0.811      0.418      -0.407      0.980
Professions intermédiaires                                   0.0677      0.363
0.186      0.852      -0.644      0.779
Retraités                                                    0.1827      0.213
0.856      0.392      -0.236      0.601
Au chômage                                                  -0.1406      0.355      -
0.396      0.692      -0.837      0.556
Etudiant, en formation ou stagiaire                         -1.3373      0.501      -
2.670      0.008      -2.319     -0.355
Occupe un emploi                                            -0.8808      0.348      -
2.530      0.011      -1.563     -0.198
Retraité(e)                                                 -0.8541      0.195      -
4.371      0.000      -1.237     -0.471
Bac ou équivalent                                           -0.1733      0.374      -
0.463      0.643      -0.907      0.560
CAP-DNB ou équivalent                                       -0.1000      0.369      -
0.271      0.787      -0.824      0.624
DUT ou équivalent                                           -0.5612      0.385      -
1.457      0.145      -1.316      0.194
Licence ou équivalent                                       -0.5905      0.391      -
1.509      0.131      -1.358      0.177
Master ou équivalent                                        -0.9696      0.448      -
2.166      0.030      -1.847     -0.092
Sans diplôme                                                 0.0736      0.371
0.199      0.843      -0.653      0.800
Zone à faible densité de population                          0.2045      0.104
1.974      0.048       0.001      0.408
Zone à forte densité de population                          -0.1072      0.106      -
1.014      0.310      -0.314      0.100
==================================================================================
=======================================
```