



Escuela Técnica Superior de Ingeniería (ICAI)

Determinantes de la carga energética excesiva en los hogares españoles a partir de microdatos de la EPF 2016–2024

Autor: Gonzalo Alonso Lidón

Director: Manuel Alejandro Betancourt Odio

MADRID | Abril 2026

Resumen

Este Trabajo Fin de Grado usa microdatos de la Encuesta de Presupuestos Familiares (EPF) del INE, periodo 2016–2024, con 183.477 hogares. Estima un logit ponderado cuya variable dependiente vale uno cuando el cociente entre el gasto energético del hogar (COICOP 04.5) y su gasto total supera el doble de la mediana ponderada. Los pesos se normalizan a la N muestral para no inflar la N efectiva del modelo, y los errores estándar son cluster-robust por comunidad autónoma \times año.

La investigación española sobre pobreza energética se apoya sobre todo en indicadores agregados o en umbrales de renta. Este trabajo aporta evidencia de qué rasgos del hogar elevan la probabilidad de destinar una proporción excesiva del presupuesto a energía.

El nivel de gasto total es el predictor más potente: los hogares del quintil más bajo tienen una probabilidad 36,5 p.p. mayor de presentar carga energética excesiva que los del quintil más alto (OR = 13,230). Siguen el alquiler inferior al mercado (AME = 12,2 p.p.; OR = 2,673), los edificios pequeños de menos de diez viviendas (8,2 p.p.), las zonas poco pobladas (6,4 p.p.) y la búsqueda de primer empleo del sustentador principal (4,9 p.p.). El modelo da un AUC de 0,778 y un pseudo- R^2 de McFadden de 0,145; la prevalencia de carga excesiva es del 12,5 %.

Desde el lado asegurador, los resultados abren aplicaciones concretas: segmentación de primas de hogar por estructura de gasto, producto paramétrico frente a picos de factura y financiación de rehabilitación vinculada a la póliza. En política pública, apuntan a intervenciones que trabajen sobre la estructura interna del presupuesto del hogar antes que sobre el nivel agregado de renta.

Palabras clave: pobreza energética, carga energética, determinantes del hogar, microdatos

EPF, España, modelo logístico, efectos marginales promedio, sector asegurador.

Abstract

This Bachelor's Thesis draws on microdata from Spain's Household Budget Survey (Encuesta de Presupuestos Familiares, EPF), published by the National Statistics Institute (INE), covering the period 2016–2024 with 183.477 households. A weighted logit estimates a binary outcome that equals one when the ratio of household energy expenditure (COICOP 04.5) to total expenditure exceeds twice the weighted median. Sampling weights are normalised to the sample N so the model's effective sample size is not inflated, and standard errors are cluster-robust by autonomous community \times year.

Spanish research on energy poverty has mostly worked with aggregate indicators or income thresholds. This thesis contributes evidence on which household features raise the probability of spending an excessive share of the budget on energy.

Total expenditure is the single strongest predictor: households in the lowest quintile are 36,5 p.p. more likely to face excessive energy burden than those in the top quintile (OR = 13,230). Next come below-market rental tenure (AME = 12,2 p.p.; OR = 2,673), buildings with fewer than ten dwellings (8,2 p.p.), low-density areas (6,4 p.p.) and a main earner searching for a first job (4,9 p.p.). The model yields an AUC of 0,778 and a McFadden pseudo- R^2 of 0,145; the prevalence of excessive burden stands at 12,5 %.

On the insurance side, the results open concrete applications: segmentation of home-insurance premiums by expenditure structure, a parametric product that cushions energy-bill spikes, and policy-linked financing of energy retrofits. On the public policy side, they argue for interventions that target the internal structure of the household budget rather than aggregate income alone.

Keywords: energy poverty, energy burden, household determinants, EPF microdata, Spain,

logistic regression, average marginal effects, insurance sector.

Índice general

Resumen	I
Abstract	III
1. Introducción	1
1.1. Motivación y contexto	1
1.2. Objetivos del trabajo	2
1.3. Estructura del documento	3
2. Revisión de la literatura	4
2.1. Definición y medición de la pobreza energética	4
2.1.1. El umbral del 10 %: la propuesta de Boardman	4
2.1.2. La revisión de Hills y el indicador LIHC	5
2.1.3. El debate europeo: indicadores y comparabilidad	6
2.1.4. El gasto energético como proporción del presupuesto	7
2.2. Determinantes de la carga energética en los hogares	7
2.2.1. Factores estructurales y el marco de Bouzarovski	8
2.2.2. La perspectiva bottom-up	8
2.2.3. Consecuencias sobre la salud	9
2.2.4. Evidencia empírica en España	9

2.2.5. Marco regulatorio y políticas públicas	10
2.2.6. Contribución de este trabajo	11
3. Metodología	12
3.1. Estrategia empírica	12
3.1.1. Efectos marginales medios (AME)	13
3.1.2. Sensibilidad del umbral de clasificación	14
3.2. Aspectos metodológicos transversales	14
3.2.1. Diseño muestral y pesos de encuesta	14
3.2.2. Diagnósticos de ajuste	15
3.2.3. Limitaciones	16
3.3. Software y reproducibilidad	16
4. Aplicación de la metodología propuesta	17
4.1. Datos: la Encuesta de Presupuestos Familiares	17
4.1.1. Periodo y muestra	17
4.1.2. Variable dependiente	18
4.1.3. Variables independientes	18
4.1.4. Nota sobre la clasificación COICOP	20
4.2. Análisis descriptivo micro	20
4.2.1. Estadística descriptiva de la muestra	20
4.2.2. Crowding-out: composición del gasto por categoría de carga	21
4.3. Modelo logístico ponderado	22
4.3.1. Especificación del modelo	22
4.3.2. Resultados	23
4.3.3. Diagnósticos de ajuste	26

4.3.4. Análisis de sensibilidad	28
4.3.5. Perfil de hogar vulnerable	28
5. Conclusiones	31
5.1. Principales hallazgos	31
5.2. Implicaciones para el sector asegurador	32
5.2.1. Segmentación de primas de hogar	33
5.2.2. Producto contra picos de gasto energético	33
5.2.3. Financiación de rehabilitación energética	34
5.3. Limitaciones y líneas futuras	34
Bibliografía	36
A. Declaración de Uso de Herramientas de Inteligencia Artificial Generativa en Trabajos Fin de Grado	39

Índice de figuras

4.1. Distribución del ENERGY_SHARE a nivel de hogar individual	21
4.2. Composición del gasto por categoría de carga energética	22
4.3. Curva ROC del modelo logístico ponderado. $AUC = 0,778$. La diagonal representa el clasificador aleatorio. Fuente: EPF microdatos INE, 2016– 2024. Elaboración propia.	26

Índice de cuadros

4.1. Variables del modelo logístico ponderado. Cada variable se clasifica por su tipo: variable dependiente (VD), predictor categórico (dummies) o predictor continuo. La columna “Categorías” indica el número de niveles; la categoría de referencia es la de mayor peso muestral ($\text{sum}(\text{FACTOR})$), salvo para los quintiles de gasto (referencia: quintil 5, mayor gasto) y el año (referencia: 2024).	19
4.2. Estadísticos descriptivos ponderados de la carga energética (gasto energía COICOP 04.5 sobre gasto total). Variable dependiente: $Y = 1$ si carga energética $\geq 2 \times$ mediana ponderada (umbral 2M). Fuente: microdatos EPF (INE), 2016–2024. Elaboración propia.	20
4.3. Resultados del modelo logístico ponderado ($2 \times$ mediana)	24
4.4. Factor de inflación de la varianza (VIF)	27
4.5. Análisis de sensibilidad del modelo: OR por umbral de carga energética (errores estándar cluster-robust por CCAA \times año)	29

Capítulo 1

Introducción

1.1. Motivación y contexto

En los últimos años, el debate sobre la pobreza energética ha cobrado fuerza en España y en el conjunto de la Unión Europea. La escalada de precios que comenzó en el otoño de 2021 y alcanzó su máximo en verano de 2022, con incrementos superiores al 40 % en el precio de la electricidad en España, expuso la fragilidad financiera de amplios segmentos de la población frente al coste de la energía (Eurostat, 2026). Aunque la introducción de la excepción ibérica y las rebajas fiscales atenuaron parcialmente el impacto, los precios se han estabilizado en niveles superiores a los previos a 2021, y el problema de fondo no ha desaparecido: hay hogares para los que la factura energética supone un esfuerzo desproporcionado dentro de su presupuesto.

Este trabajo parte de una pregunta directa: qué características del hogar determinan que la proporción del gasto energético sobre el gasto total (lo que la literatura denomina carga energética, o *energy burden*) alcance niveles excesivos. Dicho de otro modo, dados dos hogares en la misma comunidad autónoma y en el mismo año, expuestos a precios y condiciones climáticas similares, el objetivo es identificar qué factores internos del presupuesto, de la composición del hogar y de la vivienda explican que uno cruce el umbral de vulnerabilidad y el otro no.

La mayor parte de los estudios sobre pobreza energética en España se ha centrado en

diagnósticos agregados o territoriales, a partir de indicadores subjetivos de la Encuesta de Condiciones de Vida o de umbrales basados en la proporción del gasto energético sobre la renta. Son menos los trabajos que han analizado los determinantes a nivel de hogar individual con microdatos de gasto detallado. Este trabajo se sitúa en ese hueco: toma como medida principal la carga energética, definida como el cociente entre el gasto en energía (COICOP 04.5) y el gasto total del hogar (GASTOT), con microdatos de la Encuesta de Presupuestos Familiares (EPF) del INE para el periodo 2016–2024 (nueve oleadas anuales). El umbral principal es dos veces la mediana ponderada de la carga energética (indicador 2M), con análisis de sensibilidad en los umbrales del 10 % y del percentil 75. La estrategia empírica se inspira en Legendre y Ricci (2015), que estimaron un modelo logístico sobre la encuesta de hogares francesa para identificar qué hogares resultan más vulnerables bajo distintas definiciones del problema; se adapta aquí el mismo planteamiento al caso español.

El análisis se sitúa en un contexto energético que conviene tener presente. España experimentó durante el periodo de estudio una transformación notable de su mercado eléctrico: la crisis de precios de 2021–2022, la respuesta regulatoria (excepción ibérica, reducción del IVA, suspensión del IVPEE) y la posterior estabilización en niveles superiores a los históricos. El parque residencial español, con una antigüedad media elevada y una penetración desigual de las rehabilitaciones energéticas, añade una dimensión estructural al problema.

1.2. Objetivos del trabajo

El trabajo persigue dos objetivos concretos:

1. Identificar qué características del hogar (nivel de gasto total, composición familiar, régimen de tenencia, tipo de edificio, situación laboral del sustentador principal, densidad del municipio y comunidad autónoma) se asocian con una probabilidad significativamente mayor de presentar una carga energética excesiva, mediante un modelo logístico ponderado con errores estándar cluster-robust por comunidad autónoma \times año estimado sobre 183.477 observaciones de la EPF (2016–2024).

2. Verificar la robustez de los resultados ante distintas definiciones del umbral de vulnerabilidad ($2 \times$ mediana, 10 % fijo y percentil 75), de modo que las conclusiones no dependan de una convención normativa concreta.

1.3. Estructura del documento

El documento se organiza en cinco capítulos. El capítulo 2 revisa el concepto de pobreza energética, los principales indicadores de medición y la evidencia empírica sobre los determinantes de la carga energética a nivel de hogar. Sobre esa base, el capítulo 3 desarrolla la estrategia empírica: el modelo logístico ponderado, la definición del umbral y los diagnósticos de ajuste. La parte aplicada se concentra en el capítulo 4, que describe los datos de la EPF, presenta el análisis descriptivo de la muestra y estima el modelo logístico con sus resultados y análisis de sensibilidad. El capítulo 5 cierra con la síntesis de los hallazgos, sus implicaciones en el sector asegurador, y las líneas futuras de investigación.

Capítulo 2

Revisión de la literatura

2.1. Definición y medición de la pobreza energética

La pobreza energética es un concepto que, a pesar de llevar más de tres décadas en la literatura académica, todavía no cuenta con una definición universalmente aceptada. Lo que sí está claro es que se refiere a una situación en la que un hogar no puede permitirse un nivel adecuado de servicios energéticos (sobre todo calefacción, refrigeración e iluminación) sin que ello suponga un esfuerzo desproporcionado dentro de su presupuesto. Dicho de forma más directa: un hogar en pobreza energética es aquel que tiene que elegir entre pasar frío y llegar a fin de mes.

El debate académico se ha centrado, en buena medida, en *cómo medir* esa situación, y las respuestas han sido bastante distintas según la tradición y el país de origen. En esta sección se recorren las principales propuestas, desde el umbral del 10 % de Boardman (1991) hasta los enfoques más recientes que intentan capturar la naturaleza multidimensional del problema.

2.1.1. El umbral del 10 %: la propuesta de Boardman

El punto de partida del campo es el trabajo de Boardman (1991), que definió la pobreza energética en el contexto del Reino Unido como la situación en la que un hogar necesita

destinar más del 10 % de sus ingresos a mantener un nivel adecuado de calefacción. El indicador que se conoce en la literatura como **2M** (dos veces la mediana del gasto energético relativo) tiene su origen conceptual en esta propuesta, aunque el umbral fijo del 10 % fue el que realmente marcó la agenda política durante los años noventa y la primera década de los 2000.

Este enfoque tiene la ventaja de ser sencillo y fácil de comunicar, pero también presenta limitaciones importantes. Moore (2012) señala que el umbral del 10 % es arbitrario y que distintas definiciones del problema conducen a políticas muy diferentes. Además, un indicador basado exclusivamente en la proporción de gasto no distingue entre hogares que gastan mucho porque su vivienda es ineficiente y hogares que gastan mucho simplemente porque tienen ingresos muy bajos. Boardman (2010) reconoció estas limitaciones en un trabajo posterior y abogó por complementar el indicador con información sobre la eficiencia energética de la vivienda.

2.1.2. La revisión de Hills y el indicador LIHC

En 2012, el gobierno británico encargó a John Hills una revisión en profundidad del sistema de medición de la pobreza energética. El resultado fue el informe *Getting the Measure of Fuel Poverty* (Hills, 2012), que propuso el indicador **LIHC** (*Low Income High Costs*). A diferencia del umbral del 10 %, el LIHC exige que se cumplan dos condiciones simultáneas: que los costes energéticos del hogar estén por encima de la mediana nacional, y que la renta disponible del hogar, una vez descontados esos costes, quede por debajo del umbral de pobreza relativa.

Lo más relevante del informe de Hills para este trabajo no es tanto el indicador compuesto en sí, sino el concepto de *residual income* (*renta residual*) que introduce como componente del LIHC: lo que le queda al hogar después de pagar la vivienda y la energía. Esto desplaza el foco desde el ratio simple de gasto sobre renta hacia la *estructura* del presupuesto doméstico, es decir, hacia cómo se reparte el gasto del hogar entre partidas más o menos flexibles. Esa idea es precisamente la que lleva al enfoque de este TFG, aunque la forma concreta de plasmarlo en un modelo sea diferente, como se explica más adelante.

2.1.3. El debate europeo: indicadores y comparabilidad

A escala europea, la medición de la pobreza energética ha seguido un camino distinto al del Reino Unido. La fuente de datos principal ha sido la encuesta EU-SILC (*Statistics on Income and Living Conditions*), que incluye tres indicadores subjetivos ampliamente utilizados: la incapacidad de mantener la vivienda adecuadamente caliente, los retrasos en el pago de facturas y la presencia de humedades o filtraciones en la vivienda (Energy Poverty Advisory Hub, 2023; Eurostat, 2026).

Thomson y Snell (2013) fueron de los primeros en cuantificar el efecto de la pobreza energética a escala de la Unión Europea utilizando estos indicadores, y pusieron de manifiesto las enormes diferencias entre países del norte y del sur de Europa. Años más tarde, Thomson et al. (2017) plantearon la necesidad de repensar la medición en el contexto europeo, argumentando que los indicadores existentes, tanto los basados en gasto como los subjetivos, capturan aspectos distintos del problema y que ninguno de ellos es suficiente por sí solo.

Dubois y Meier (2016) introdujeron una distinción útil entre *energy affordability* (asequibilidad energética) y *energy poverty* (pobreza energética), señalando que la primera se refiere a la capacidad de pago en un momento dado, mientras que la segunda implica una situación más estructural y persistente. Esta distinción es relevante porque el indicador que se utiliza en este trabajo (la proporción del gasto energético sobre el gasto total) es, estrictamente hablando, un indicador de asequibilidad (captura el esfuerzo en un momento dado), aunque cuando se observa de forma consistente a lo largo de varios años puede interpretarse como un indicador de pobreza energética en un sentido más amplio.

Otros trabajos han ampliado la perspectiva de medición. Pelz et al. (2018) ofrecen una revisión crítica de los enfoques multidimensionales, mientras que Okushima (2016) aplica indicadores basados en gasto al caso japonés, mostrando que el problema no es exclusivo de Europa. Maxim et al. (2016) analizan las implicaciones de utilizar distintos indicadores a escala de la UE, y sus resultados confirman que la elección del indicador condiciona de forma significativa la identificación de los hogares vulnerables.

2.1.4. El gasto energético como proporción del presupuesto

El indicador que se utiliza en este trabajo es la **carga energética** (*energy burden*): la proporción del gasto total del hogar que se destina a energía. Formalmente:

$$\text{Carga energética} = \frac{\text{Gasto en energía (COICOP 04.5)}}{\text{Gasto total del hogar (GASTOT)}} \quad (2.1)$$

La elección de este indicador responde a una razón metodológica concreta: la Encuesta de Presupuestos Familiares (EPF) del INE, que constituye la fuente de datos de este trabajo, no contiene información directa sobre la renta disponible de los hogares. La EPF recoge con detalle el gasto anual del hogar desagregado por categorías COICOP, pero no incluye variables de ingreso. Por tanto, el denominador del indicador es el gasto total (GASTOT), no la renta.

Esta elección es coherente con la literatura. Legendre y Ricci (2015) utilizan un enfoque basado en el gasto para medir la vulnerabilidad energética con la encuesta de hogares francesa, argumentando que el gasto observado refleja mejor el comportamiento real del hogar que la renta declarada, que puede estar sujeta a fluctuaciones transitorias. Deaton y Muellbauer (1980) establecieron además que, en la teoría del consumo, el gasto total es un mejor proxy del ingreso permanente del hogar que la renta corriente, lo que refuerza la validez del denominador utilizado.

El umbral principal de este trabajo es **2M** (dos veces la mediana ponderada de la carga energética), que en la muestra EPF 2016–2024 se sitúa en torno al 7,8 %. Como análisis de sensibilidad se utilizan también el umbral fijo del 10 % y el percentil 75, siguiendo la práctica habitual en la literatura (Romero et al., 2018).

2.2. Determinantes de la carga energética en los hogares

La segunda pregunta que plantea la literatura es: una vez que se ha definido y medido la pobreza energética, *qué factores explican* que unos hogares presenten una carga energética excesiva y otros no. Este bloque recoge las principales aportaciones sobre los

determinantes de la vulnerabilidad energética a nivel de hogar.

2.2.1. Factores estructurales y el marco de Bouzarovski

Bouzarovski y Petrova (2015) propusieron un marco conceptual que entiende la pobreza energética como el resultado de la interacción entre tres factores: unos ingresos bajos, unos precios energéticos altos y una vivienda ineficiente desde el punto de vista energético. Este enfoque sitúa el problema fuera de la insuficiencia de renta e incorpora la dimensión física de la vivienda como factor determinante.

Bouzarovski (2018) amplió este marco en un trabajo posterior, argumentando que la pobreza energética en Europa está muy ligada a lo que denomina la *brecha infraestructural* del continente: las enormes diferencias en la calidad del parque residencial entre países y dentro de cada país. Según esta perspectiva, la vulnerabilidad energética no es solo una cuestión de renta, sino en muchos de los casos, de la calidad del inmueble.

2.2.2. La perspectiva bottom-up

Frente a los enfoques que analizan la pobreza energética desde arriba (indicadores agregados, estadísticas nacionales), una línea de investigación ha intentado entender el problema desde la experiencia de los hogares afectados.

Middlemiss y Gillard (2015) entrevistaron a hogares en situación de pobreza energética en el Reino Unido y mostraron que la capacidad de un hogar para afrontar un incremento en los precios de la energía no depende solo de su nivel de renta, sino también de cómo está organizado su presupuesto. Así, hogares con ingresos parecidos pueden presentar niveles de vulnerabilidad muy distintos según el peso de sus gastos comprometidos (alquiler, suministros básicos, deuda). Middlemiss (2017) amplió esta perspectiva con un análisis de las políticas de pobreza energética en Inglaterra, argumentando que la respuesta política ha tendido a individualizar el problema (responsabilizando al hogar de su consumo) cuando en realidad se trata de un problema estructural.

Tirado Herrero y Ürge-Vorsatz (2012) documentaron el fenómeno de la trampa energéti-

ca en países centroeuropeos: viviendas con muy baja eficiencia energética heredadas del periodo comunista obligan a los hogares a destinar una proporción desmedida de su presupuesto a calefacción, sin posibilidad real de invertir en mejoras. Healy (2004) mostró que la calidad de la vivienda es un determinante clave de la pobreza energética en toda Europa, y que los efectos sobre la salud de vivir en viviendas frías o húmedas son significativos y medibles.

2.2.3. Consecuencias sobre la salud

Liddell y Morris (2010) realizaron una revisión de la evidencia disponible sobre la relación entre pobreza energética y salud, concluyendo que vivir en una vivienda fría o energéticamente inadecuada se asocia con peores resultados de salud respiratoria, cardiovascular y mental. Hernández (2016) amplió esta perspectiva al caso estadounidense a través del concepto de *energy insecurity* (inseguridad energética): la ausencia de un margen financiero mínimo puede hacer que un retraso en el pago de la factura energética termine convirtiéndose en un problema de salud, al forzar al hogar a elegir entre calefacción o alimentación.

Estas consecuencias sobre la salud refuerzan la relevancia del análisis de los determinantes: identificar qué características del hogar predicen una carga energética excesiva es un paso necesario para diseñar intervenciones que reduzcan tanto la vulnerabilidad económica como sus efectos sobre el bienestar.

2.2.4. Evidencia empírica en España

En España, la investigación sobre pobreza energética ha crecido de forma notable durante la última década, impulsada en parte por la crisis de precios energéticos de 2021–2022. Sin embargo, buena parte de los estudios se han centrado en diagnósticos agregados o territoriales, y son menos los que han analizado los determinantes a nivel de hogar con microdatos.

Tirado Herrero et al. (2018) elaboraron un informe de referencia para la Asociación de Ciencias Ambientales (ACA) en el que aplican al caso español un sistema de indicadores

de pobreza energética, incluyendo el indicador 2M y la incapacidad de mantener la vivienda adecuadamente caliente. Este trabajo constituye uno de los diagnósticos más completos del problema en España y ofrece datos desagregados por comunidades autónomas.

Romero et al. (2014) realizaron uno de los primeros análisis económicos de la pobreza energética en España, cuantificando su incidencia y proponiendo medidas de actuación. Luego relacionaron la medición del problema con el diseño de políticas públicas, mostrando que la elección del indicador condiciona significativamente tanto la identificación de los hogares vulnerables como las políticas que se derivan.

Sánchez-Guevara Sánchez et al. (2019) analizaron la pobreza energética en Madrid utilizando un enfoque de confort adaptativo, que tiene en cuenta las expectativas y hábitos de los hogares, no solo el gasto objetivo. Costa-Campi et al. (2019) estudiaron las tendencias, determinantes e implicaciones de política de la pobreza energética en España utilizando datos de la EPF, lo que convierte su trabajo en uno de los antecedentes más directos de este TFG.

2.2.5. Marco regulatorio y políticas públicas

En España, la respuesta política a la pobreza energética se apoya en tres ejes principales: protección directa al consumidor, regulación de precios y rehabilitación del parque inmobiliario.

El principal instrumento de protección es el **Bono Social Eléctrico**, un descuento del 25 %–40 % en la factura de electricidad para hogares vulnerables o en riesgo de exclusión social. Desde 2022 se amplió a hogares afectados por la crisis de precios. El **Bono Social Térmico** proporciona además una compensación anual para gastos de calefacción.

La crisis energética de 2021–2023 desencadenó un paquete de medidas de emergencia sin precedentes: reducción del IVA sobre la electricidad del 21 % al 5 %, suspensión temporal del Impuesto sobre el Valor de la Producción de Energía Eléctrica (IVPEE), y la llamada *excepción ibérica*, un límite al precio del gas natural en la generación eléctrica entre junio de 2022 y diciembre de 2023, que redujo el precio mayorista en torno a un 15 %–20 % respecto al resto del mercado europeo.

A escala estratégica, la **Estrategia Nacional contra la Pobreza Energética 2019–2024** (Ministerio para la Transición Ecológica y el Reto Demográfico, 2019) estableció el marco general de seguimiento del problema, mientras que el Plan de Recuperación, Transformación y Resiliencia (PRTR) destina alrededor de 6.800 millones de euros a la rehabilitación energética de edificios. La evolución de la carga energética observada entre 2022 y 2024 debe interpretarse teniendo en cuenta estas medidas ya que parte de la reducción observada responde a la retirada progresiva de las ayudas fiscales de emergencia, más que a una mejora estructural de la eficiencia o de la renta.

2.2.6. Contribución de este trabajo

Frente a la literatura sobre pobreza energética en España, que tiende a centrarse en indicadores agregados o umbrales basados en renta, este trabajo aporta un análisis de los determinantes a nivel de hogar utilizando la proporción del gasto energético sobre el gasto total como medida principal.

En concreto, la estrategia empírica se inspira metodológicamente en Legendre y Ricci (2015), que estimaron un modelo logístico con la encuesta de hogares francesa para identificar qué hogares resultan más vulnerables bajo distintas definiciones del problema. Se adapta ese planteamiento a los microdatos de la EPF del INE, cubriendo nueve años (2016–2024) y utilizando la carga energética (*energy burden*) como variable dependiente del modelo.

La pregunta de investigación que guía el análisis es directa: *qué características del hogar (nivel de gasto, composición familiar, régimen de tenencia, comunidad autónoma de residencia) se asocian con una probabilidad significativamente mayor de presentar una carga energética excesiva*. El capítulo de metodología (capítulo 3) detalla el procedimiento para abordar esta pregunta, y el capítulo de aplicación (capítulo 4) presenta los resultados del modelo estimado.

Capítulo 3

Metodología

3.1. Estrategia empírica

El modelo principal es una regresión logística ponderada con variable dependiente binaria. La variable toma valor 1 si la carga energética del hogar (definida como el cociente entre el gasto en energía COICOP 04.5 y el gasto total, GASTOT) supera el umbral de dos veces la mediana ponderada (indicador 2M), y valor 0 en caso contrario. En la muestra EPF 2016–2024, este umbral se sitúa en torno al 7,8 %, lo que clasifica al 12,5 % de los hogares como vulnerables.

El denominador del indicador es el gasto total, no la renta: la EPF no recoge el ingreso disponible del hogar de forma sistemática y, como se discute en el marco teórico (sección [2.1.4](#)), el gasto total es un mejor proxy del ingreso permanente que la renta corriente.

La especificación estimada es:

$$\begin{aligned}
\text{logit}(P[Y = 1]) = & \alpha + \sum_{q=1}^4 \beta_q \cdot \mathbb{1}[\text{gastot_}q = q] \\
& + \delta'_{\text{regten}} \mathbf{D}_{\text{regten}} + \delta'_{\text{tiphogar}} \mathbf{D}_{\text{tiphogar}} + \delta'_{\text{tipoedif}} \mathbf{D}_{\text{tipoedif}} \\
& + \delta'_{\text{situactsp}} \mathbf{D}_{\text{situactsp}} + \delta'_{\text{densi}} \mathbf{D}_{\text{densi}} \\
& + \beta_{\text{annocon}} \cdot \text{ANNOCON} + \delta'_{\text{tipocasa}} \mathbf{D}_{\text{tipocasa}} \\
& + \gamma' \mathbf{D}_{\text{ccaa}} + \tau' \mathbf{D}_{\text{year}} + \varepsilon
\end{aligned} \tag{3.1}$$

donde los quintiles de gasto total ($\text{gastot_}q$) se calculan por año con referencia en el quintil más alto ($q = 5$). Los bloques de dummies controlan por régimen de tenencia (REGTEN, 5 categorías), composición del hogar (TIPHOGAR1, 12 categorías), tipo de edificio (TIPOEDIF, 5 categorías), situación laboral del sustentador principal (SITUACTSP, 8 categorías), densidad del municipio (DENSIDAD, 3 categorías), tipo de vivienda (TIPOCASA) y antigüedad de la construcción (ANNOCON, variable continua). \mathbf{D}_{ccaa} son 17 dummies de comunidad autónoma (Ceuta y Melilla combinadas) y \mathbf{D}_{year} son 8 dummies de año (referencia: 2024), introducidas al apilar los nueve años de la EPF para absorber diferencias temporales.

La variabilidad climática entre territorios se absorbe por los efectos fijos de comunidad autónoma: la zona climática del Código Técnico de la Edificación no se introduce como variable separada porque está altamente correlacionada con las dummies autonómicas en España. Los errores estándar son cluster-robust por comunidad autónoma \times año. Los resultados se reportan como *odds ratios* (OR) con intervalos de confianza al 95 % y como efectos marginales medios (*average marginal effects*, AME).

3.1.1. Efectos marginales medios (AME)

Los *odds ratios* miden la variación relativa de la variable dependiente, pero su magnitud no es interpretable como un efecto sobre la probabilidad. Para facilitar la lectura de los resultados, se reportan junto a los OR los efectos marginales medios, calculados como:

$$\widehat{AME}_k = \frac{1}{N} \sum_{i=1}^N \frac{\partial \hat{P}(Y = 1 | \mathbf{x}_i)}{\partial x_{k,i}} \quad (3.2)$$

En la práctica, el cálculo se realiza mediante diferencias finitas: para cada observación, se evalúa la probabilidad predicha con y sin el cambio unitario en la variable de interés, y la diferencia se promedia sobre la muestra ponderada. Esta aproximación es necesaria porque la implementación GLM de `statsmodels` no dispone de un método `get_margeff` compatible con pesos de frecuencia.

3.1.2. Sensibilidad del umbral de clasificación

Para verificar la estabilidad de los resultados ante la definición de la variable dependiente, se reestima el modelo logístico con tres umbrales:

- $Y = 1$ si la carga energética $\geq 10\%$ (umbral fijo de Hills/Boardman, transpuesto al eje gasto).
- $Y = 1$ si la carga energética $\geq 2 \times$ mediana ponderada (umbral principal del trabajo; 7,8 % en la muestra).
- $Y = 1$ si la carga energética $\geq P_{75}$ muestral ($\approx 5,9\%$).

Los resultados de los tres modelos se presentan en una tabla comparativa con los odds ratios por columna. Si los signos y la jerarquía relativa de los predictores se mantienen estables, las conclusiones no dependen de la convención adoptada.

3.2. Aspectos metodológicos transversales

3.2.1. Diseño muestral y pesos de encuesta

La EPF es una encuesta con diseño muestral complejo (estratificación, conglomerados y pesos de elevación) y estructura de secciones cruzadas repetidas, cuya descripción detallada se presenta en el capítulo 4.1.

Todos los estadísticos descriptivos y los modelos se estiman utilizando los pesos de elevación (FACTOR) de la encuesta. En las regresiones, estos pesos se normalizan antes de incorporarlos al modelo, de forma que su suma coincida con el tamaño real de la muestra. Así se respeta la estructura muestral del INE sin inflar artificialmente el número de observaciones que el programa utiliza al calcular la precisión de las estimaciones.

Además, los errores estándar se calculan con una corrección *cluster-robust*, agrupando las observaciones por comunidad autónoma y año de encuesta. Esto permite tener en cuenta que los hogares de una misma región y un mismo año pueden estar afectados por factores comunes, como cambios de precios, condiciones climáticas o medidas de política energética.

Al apilar los nueve estudios (uno por año) en un análisis conjunto (*pooled cross-section*), los pesos se acumulan por año sin recalibración, y se incluyen efectos fijos de año para controlar las posibles diferencias temporales.

3.2.2. Diagnósticos de ajuste

La bondad del modelo se evalúa mediante tres indicadores complementarios:

- **AUC-ROC**: área bajo la curva ROC, que mide la capacidad discriminativa del modelo (un clasificador aleatorio obtiene 0,5; un clasificador perfecto, 1,0).
- **Pseudo- R^2 de McFadden**: cociente entre la log-verosimilitud del modelo completo y la del modelo nulo, que cuantifica la mejora en ajuste atribuible a los predictores.
- **Brier score**: media de las diferencias al cuadrado entre la probabilidad predicha y el resultado observado, con valores más bajos indicando mejor calibración.

Para evaluar la multicolinealidad entre predictores se calcula el factor de inflación de la varianza (VIF, *variance inflation factor*) de cada variable. Valores por debajo de 10 se consideran aceptables y por debajo de 5, conservadores.

3.2.3. Limitaciones

1. **Calidad de vivienda no disponible en EPF:** la EPF no recoge variables de eficiencia energética de la vivienda (calificación energética, consumo en kWh). La antigüedad de la construcción y el tipo de edificio aproximan parcialmente esta dimensión, pero no la sustituyen.
2. **Precio de la energía no individualizable:** la EPF registra el gasto en energía, no el precio unitario ni la cantidad consumida. El precio de la energía no entra al modelo como regresor; su variación temporal se absorbe por los efectos fijos de año y su variación territorial, por los efectos fijos de comunidad autónoma.
3. **Correlación, no causalidad:** los datos de secciones cruzadas repetidas de la EPF no permiten identificar efectos causales. Las variables del modelo pueden estar correlacionadas con factores no observados que también determinan la carga energética. Los resultados capturan asociaciones estadísticas, no efectos causales.

3.3. Software y reproducibilidad

El análisis se realiza en **Python 3.9** con las librerías `pandas`, `numpy`, `scipy` y `statsmodels`. La estimación del modelo logístico se ejecuta mediante `sm.GLM` con familia Binomial y pesos de frecuencia (`FACTOR`).

Capítulo 4

Aplicación de la metodología propuesta

4.1. Datos: la Encuesta de Presupuestos Familiares

La fuente de datos del trabajo es la Encuesta de Presupuestos Familiares (EPF), elaborada por el Instituto Nacional de Estadística (INE). La EPF es una encuesta anual que recoge información detallada sobre el gasto de los hogares españoles desagregado por categorías COICOP, junto con variables sociodemográficas del hogar y del sustentador principal.

El diseño muestral es de secciones cruzadas repetidas (*repeated cross-section*): cada año se encuesta a una muestra representativa de aproximadamente 22.000 hogares, con una renovación de alrededor del 50 % entre estudios consecutivos. Esto significa que el mismo hogar puede aparecer en dos años seguidos, pero la encuesta no está diseñada como un panel longitudinal: no se sigue a los mismos hogares a lo largo del tiempo.

4.1.1. Periodo y muestra

El análisis utiliza las nueve años disponibles de la EPF, correspondientes al periodo 2016–2024. Tras la limpieza de datos (eliminación de hogares con gasto energético nulo, que representan situaciones atípicas como viviendas desocupadas o errores de registro), la muestra comprende 183.477 observaciones.

Los ficheros de microdatos de uso público son de descarga libre en la web del INE.

4.1.2. Variable dependiente

La variable dependiente del modelo es una variable binaria que indica si la carga energética del hogar supera el umbral de vulnerabilidad. La carga energética se define como:

$$\text{Carga energética}_i = \frac{G_{e,i}}{G_{\text{total},i}} \quad (4.1)$$

donde $G_{e,i}$ es el gasto anual del hogar i en energía (COICOP 04.5: electricidad, gas natural y otros combustibles) y $G_{\text{total},i}$ es el gasto total del hogar (campo GASTOT de la EPF).

El umbral principal es el indicador 2M: dos veces la mediana ponderada de la carga energética, calculada con los pesos de elevación (FACTOR) de cada oleada. En la muestra EPF 2016–2024, la mediana ponderada se sitúa en el 3,9 % y el umbral 2M, en el 7,8 %. Con esta definición, el 12,5 % de los hogares se clasifica como vulnerables. Como análisis de sensibilidad se utilizan también el umbral fijo del 10 % y el percentil 75 muestral (sección 3.1.2).

4.1.3. Variables independientes

La tabla 4.1 resume las variables incluidas en el modelo logístico. Todas proceden de la EPF.

Los quintiles de gasto total (gastot_q) se calculan dentro de cada año para que la clasificación refleje la posición relativa del hogar en la distribución de gasto de su año, no de la muestra conjunta. La categoría de referencia es el quintil más alto ($q = 5$), de modo que los coeficientes del modelo miden la diferencia en la probabilidad de vulnerabilidad respecto a los hogares con mayor gasto total.

El régimen de tenencia (REGTEN) distingue entre propiedad sin hipoteca, propiedad con pagos pendientes, alquiler a precio inferior al de mercado, alquiler a precio de mercado, cesión gratuita y otras situaciones. La composición del hogar (TIPHOGAR1) recoge 12 tipologías que van desde la persona sola mayor de 65 años hasta los hogares con tres o más adultos con menores.

Cuadro 4.1: Variables del modelo logístico ponderado. Cada variable se clasifica por su tipo: variable dependiente (VD), predictor categórico (dummies) o predictor continuo. La columna “Categorías” indica el número de niveles; la categoría de referencia es la de mayor peso muestral ($\text{sum}(\text{FACTOR})$), salvo para los quintiles de gasto (referencia: quintil 5, mayor gasto) y el año (referencia: 2024).

Variable	Descripción	Tipo	Categorías
<i>Variable dependiente</i>			
Carga energética	$G_e/G_{\text{total}} > 2 \times \text{mediana}$	VD (binaria)	—
<i>Predictores categóricos (dummies)</i>			
gastot_q	Quintil de gasto total del hogar (por año)	Dummies	5 (ref: q5)
REGTEN	Régimen de tenencia de la vivienda	Dummies	6 (ref: 1)
TIPHOGAR1	Composición del hogar	Dummies	12 (ref: mayor FACTOR)
TIPOEDIF	Tipo de edificio	Dummies	5 (ref: mayor FACTOR)
SITUACTSP	Situación laboral del sustentador principal	Dummies	8 (ref: 1)
DENSIDAD	Densidad de población del municipio	Dummies	3 (ref: 1)
TIPOCASA	Tipo de vivienda	Dummies	3
CCAA	Comunidad autónoma (Ceuta y Melilla combinadas)	Dummies	18 (ref: mayor FACTOR)
AÑO	Año de la oleada EPF	Dummies	9 (ref: 2024)
<i>Predictor continuo</i>			
ANNOCON	Antigüedad de la construcción (año)	Continua	—
<i>Peso muestral</i>			
FACTOR	Factor de elevación poblacional	Peso	—

Fuente: INE, EPF, ediciones 2016–2024. La especificación completa corresponde a la ecuación (3.1) del capítulo 3. Elaboración propia.

4.1.4. Nota sobre la clasificación COICOP

A partir de 2024, el INE adopta la clasificación COICOP 2024, compatible hacia atrás con la serie COICOP 2018 revisada utilizada en este estudio para el periodo 2016–2023. La partida de gasto en energía del hogar corresponde en ambas clasificaciones al grupo 04.5 (electricidad, gas y otros combustibles), por lo que la serie es directamente comparable a lo largo de los nueve estudios sin necesidad de ajustes entre ellos.

4.2. Análisis descriptivo micro

Los factores externos (precio de la energía, parque inmobiliario, contexto territorial) explican la tendencia central de la carga energética en España, pero no explican por qué dos hogares expuestos al mismo contexto tienen cargas muy distintas. Esta sección examina la estructura interna del presupuesto del hogar a partir de los 183.477 hogares de la muestra EPF (2016–2024), antes de pasar al modelo logístico de la sección 4.3.

4.2.1. Estadística descriptiva de la muestra

La tabla 4.2 resume los principales estadísticos ponderados de la carga energética en la muestra EPF utilizada en el análisis.

Cuadro 4.2: Estadísticos descriptivos ponderados de la carga energética (gasto energía COICOP 04.5 sobre gasto total). Variable dependiente: $Y = 1$ si carga energética $\geq 2 \times$ mediana ponderada (umbral 2M). Fuente: microdatos EPF (INE), 2016–2024. Elaboración propia.

Estadístico	Valor
Carga energética (media ponderada)	0,047
Carga energética (mediana ponderada)	0,039
Carga energética (DE ponderada)	0,032
Hogares con carga excesiva, umbral 2M (%)	12,5 %
N hogares (muestra conjunta)	183.511
Periodo	2016–2024

La distribución de la carga energética presenta marcada asimetría positiva (media 4,7 % frente a mediana 3,9 %), lo que indica que la mayoría de los hogares destina un porcentaje

reducido de su gasto a energía, pero la cola derecha contiene hogares con cargas muy elevadas. La figura 4.1 muestra esta distribución. La línea discontinua naranja marca el umbral 2M ($\approx 7,8\%$), que es el umbral principal del modelo logístico; la línea punteada gris indica el percentil 75 ($\approx 5,9\%$) y la línea continua, el 10% fijo, ambos utilizados como umbrales de sensibilidad.

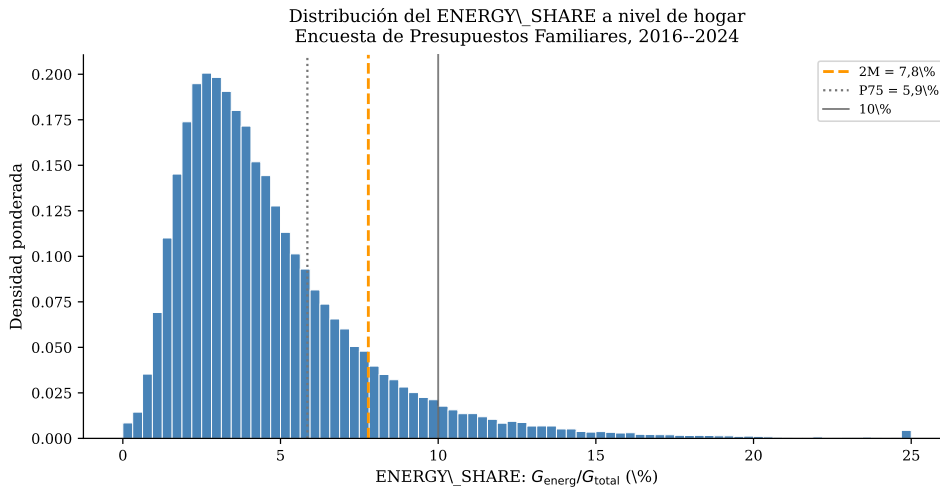


Figura 4.1: Distribución del ENERGY_SHARE (%) a nivel de hogar individual. La línea discontinua naranja marca el umbral 2M (dos veces la mediana ponderada, $\approx 7,8\%$), empleado como umbral principal en los modelos del capítulo 4.3. Se muestran además el percentil 75 ($\approx 5,9\%$) y el 10% fijo como umbrales de sensibilidad. La distribución exhibe asimetría positiva pronunciada: media 4,7% frente a mediana 3,9%. Fuente: EPF microdatos INE, 2016-2024. Elaboración propia.

Con el umbral 2M descrito en el capítulo 3 (carga energética $\geq 2 \times$ la mediana ponderada, es decir $\geq 7,8$), el 12,5% de los hogares se clasifica como vulnerable. Aunque se trate de una fracción reducida de la muestra, concentra los mecanismos de vulnerabilidad que este trabajo busca identificar.

4.2.2. Crowding-out: composición del gasto por categoría de carga

Para ilustrar el mecanismo de desplazamiento del gasto, la figura 4.2 agrupa los hogares en tres tramos descriptivos de carga energética ($<10\%$, $10-15\%$, $\geq 15\%$) y presenta la composición de su gasto por grandes partidas COICOP. Estos tramos no coinciden con el umbral 2M del modelo logístico; se utilizan aquí para visualizar cómo varía la estructura del presupuesto a medida que aumenta la proporción destinada a energía.

El indicador de margen disponible tras compromisos fijos es $s_{\text{resto},i} = (G_{\text{total},i} - G_{\text{vivienda},i} - G_{\text{energía},i}) / G_{\text{total},i}$. Este indicador no se incluye como predictor en los modelos del capítulo 4.3 (comparte el denominador con la variable dependiente, lo que introduciría correlación mecánica), sino que sirve como evidencia descriptiva del *desplazamiento del gasto*: los hogares con carga alta presentan un s_{resto} menor, lo que indica que la energía comprime el presupuesto disponible para gasto discrecional.

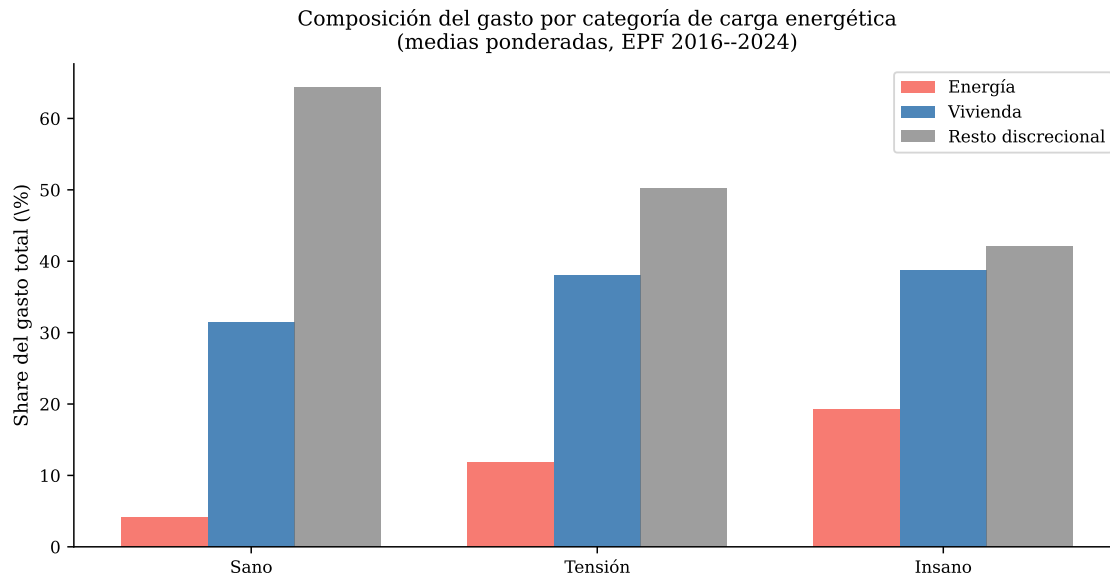


Figura 4.2: Composición del gasto por categoría de carga energética. A medida que aumenta la carga, la participación de energía y vivienda crece impactando directamente en el gasto discrecional. Fuente: EPF microdatos INE, 2016–2024. Elaboración propia.

4.3. Modelo logístico ponderado

Esta sección presenta los resultados del modelo logístico ponderado que identifica qué características del hogar se asocian con una probabilidad significativamente mayor de presentar una carga energética excesiva.

4.3.1. Especificación del modelo

La variable dependiente binaria Y toma valor 1 si la carga energética del hogar (gasto en energía COICOP 04.5 dividido por el gasto total) supera el umbral 2M, y valor 0 en caso

contrario. La especificación funcional completa fue presentada en la ecuación (3.1) del capítulo de metodología (3).

El modelo se estima mediante GLM con familia Binomial, utilizando los pesos de elevación (FACTOR) normalizados a la N muestral y errores estándar cluster-robust por comunidad autónoma \times año. Las categorías de referencia son: quintil 5 de gasto total (mayor gasto), la categoría con mayor peso muestral para las restantes variables categóricas, y el año 2024 para los efectos fijos temporales.

4.3.2. Resultados

La tabla 4.3 presenta los resultados del modelo principal, ordenados de mayor a menor efecto marginal en valor absoluto. La métrica de cabecera es el efecto marginal promedio (AME) en puntos porcentuales, que se acompaña del *odds ratio* (OR) con su intervalo de confianza al 95 % como complemento. Se comentan a continuación los predictores más relevantes.

Quintiles de gasto total. El nivel de gasto total es el predictor más potente. Los hogares del quintil más bajo tienen una probabilidad 36,5 p.p. mayor de presentar carga energética excesiva que los del quintil más alto (OR = 13,230). Los quintiles segundo (AME = 23,7 p.p.; OR = 6,463) y tercero (AME = 16,1 p.p.; OR = 3,877) también muestran efectos de gran magnitud. El resultado es aritméticamente esperable (a menor gasto total, mayor peso relativo de cualquier partida fija), pero su magnitud confirma que el gasto es el factor dominante en la probabilidad de vulnerabilidad energética.

Régimen de tenencia. Los hogares en alquiler inferior al mercado (REGTEN_4) presentan una probabilidad 12,2 p.p. mayor que la categoría de referencia (OR = 2,673), lo que los sitúa como el segundo grupo con mayor probabilidad de vulnerabilidad tras los quintiles bajos de gasto. Refleja que los hogares que acceden a rentas por debajo del mercado ya parten de una situación económica comprometida, y además la condición de inquilino limita la capacidad de invertir en mejoras de eficiencia energética del inmueble (la decisión corresponde al propietario). Los hogares en alquiler a precio de mercado (REGTEN_3) también presentan riesgo elevado, aunque algo menor (AME \approx 6,6 p.p.;

Cuadro 4.3: Resultados del modelo logístico ponderado (2× mediana)

Variable	AME (p.p.)	SE(AME)	IC 95% AME	OR	p-valor
Quintil 1 gasto	36,5***	1,01	[34,5; 38,4]	13,230	0,000
Quintil 2 gasto	23,7***	0,97	[21,8; 25,6]	6,463	0,000
Quintil 3 gasto	16,1***	0,89	[14,3; 17,8]	3,877	0,000
Alquiler inferior al mercado	12,2***	0,72	[10,8; 13,6]	2,673	0,000
Quintil 4 gasto	9,4***	0,80	[7,8; 11,0]	2,327	0,000
Edif. <10 viviendas	8,2***	0,42	[7,4; 9,0]	2,076	0,000
Alquiler a precio de mercado	6,6***	0,42	[5,8; 7,4]	1,849	0,000
Zona poco poblada	6,4***	0,38	[5,6; 7,1]	1,829	0,000
Persona sola <65	5,5***	0,43	[4,6; 6,3]	1,672	0,000
Parado (busca primer empleo)	4,9***	0,39	[4,2; 5,7]	1,582	0,000
Edif. 10+ viviendas	4,7***	0,25	[4,2; 5,2]	1,581	0,000
Jubilado	4,6***	0,37	[3,8; 5,3]	1,539	0,000
Zona intermedia	4,0***	0,29	[3,4; 4,6]	1,487	0,000
Otra situacion	3,3***	0,95	[1,5; 5,2]	1,378	0,001
Persona sola (otro)	2,9***	0,41	[2,1; 3,7]	1,329	0,000
Pareja con 1 hijo	-2,9***	0,75	[-4,4; -1,4]	0,715	0,000
Incapacitado permanente	2,8***	0,59	[1,6; 4,0]	1,313	0,000
Vivienda unifamiliar	2,7***	0,28	[2,1; 3,2]	1,305	0,000
Pareja con 2 hijos	2,6***	0,35	[1,9; 3,3]	1,297	0,000
Pareja sin hijos ambos 65+	-2,5***	0,48	[-3,4; -1,5]	0,755	0,000
Ocupado por cuenta ajena temp.	2,4***	0,28	[1,9; 3,0]	1,280	0,000
Otro tipo de hogar (2 adultos)	-2,4***	0,98	[-4,3; -0,4]	0,766	0,009
Otro tipo de hogar (3+ adultos)	2,1***	0,40	[1,3; 2,9]	1,236	0,000
Pareja con 3+ hijos	-1,7***	0,47	[-2,7; -0,8]	0,825	0,000
Ocupado por cuenta ajena indef.	1,6**	0,69	[0,3; 3,0]	1,179	0,023
Adulto con hijos	-1,6***	0,47	[-2,5; -0,7]	0,838	0,000
Persona sola 65+	1,4***	0,43	[0,6; 2,3]	1,157	0,001
Otra forma de tenencia	1,2	0,75	[-0,3; 2,7]	1,129	0,120
Pareja sin hijos	1,0	1,58	[-2,1; 4,1]	1,111	0,528
Cesion gratuita	1,0	0,64	[-0,3; 2,3]	1,108	0,129
Parado (ha trabajado)	0,9	5,03	[-9,0; 10,7]	1,092	0,868
Propiedad con hipoteca	0,6**	0,26	[0,1; 1,1]	1,062	0,025
Piso	0,3	0,37	[-0,4; 1,0]	1,032	0,416
Otro tipo de edificio	0,2	2,73	[-5,2; 5,6]	1,021	0,943
Antes de 1980	-0,2	0,31	[-0,8; 0,5]	0,984	0,616
Vivienda independiente	0,0	0,44	[-0,8; 0,9]	1,004	0,935

Nota: AME en puntos porcentuales (p.p.) sobre la probabilidad. Efectos fijos de CCAA incluidos (17 dummies, ref: 1). Efectos fijos temporales incluidos (8 dummies, ref: 2024). *** p<0,01; ** p<0,05; * p<0,1. Errores estándar cluster-robust por CCAA×año; pesos FACTOR normalizados.

OR = 1,849).

Tipo de edificio. Los hogares en edificios pequeños, de menos de diez viviendas (TIPOEDIF_1), elevan la probabilidad de carga energética excesiva en 8,2 p.p. respecto a la categoría de referencia (OR = 2,076). Lo más probable al tener más superficie que calentar por vivienda que los bloques grandes hay una mayor demanda de calefacción y refrigeración.

Composición del hogar. Los hogares unipersonales de menos de 65 años (TIPHOGAR1_1) constituyen un grupo especialmente vulnerable: su probabilidad de superar el umbral es 5,5 p.p. mayor que la de la categoría de referencia (OR = 1,672). El patrón encaja con un consumo energético relativamente fijo para un solo ocupante que no se beneficia de las economías de escala de un hogar mayor.

Situación laboral. Los hogares cuyo sustentador principal busca un primer empleo (SITUACTSP_6) presentan una probabilidad 4,9 p.p. mayor de superar el umbral, y los hogares con sustentador jubilado (SITUACTSP_3) una probabilidad 4,6 p.p. mayor. Ambos grupos son significativamente más vulnerables que los hogares con sustentador ocupado, lo que apunta a la menor capacidad de absorber subidas de factura cuando el ingreso laboral es pequeño o nulo.

Densidad del municipio. Los hogares en municipios de baja densidad (DENSIDAD_3) tienen una probabilidad 6,4 p.p. mayor de vulnerabilidad que los de zonas densas (OR = 1,829). Al vivir en entornos rurales, la dependencia en combustibles no de red y las menores opciones de transporte público obligan a que el gasto energético sea más elevado.

Antigüedad de la construcción. La variable ANNOCON se codifica como dummy para viviendas anteriores a 1980 (ANNOCON_1) frente a viviendas de 1980 en adelante (categoría de referencia). El AME estimado es de apenas 0,16 p.p. y no es estadísticamente significativo ($p = 0,616$; OR = 0,984), de manera que, una vez controlado el resto de factores, la antigüedad de la vivienda no tiene un efecto diferencial sobre la probabilidad de carga excesiva en este estudio.

Comunidad autónoma. Los efectos fijos de comunidad autónoma capturan las diferencias territoriales no observadas (clima, precios regulados, composición del parque residencial). Destaca Ceuta y Melilla (CCAA_combined_13), con un AME de 9,8 p.p. re-

specto a la CCAA de referencia (Andalucía, $OR = 2,364$), lo que refleja las condiciones específicas de estos territorios.

4.3.3. Diagnósticos de ajuste

El modelo presenta un AUC-ROC de 0,778, lo que indica una capacidad de predicción aceptable; el modelo distingue razonablemente bien entre hogares que superan el umbral y hogares que no lo superan. El pseudo- R^2 de McFadden es 0,145 y el Brier score es 0,100.

La figura 4.3 muestra la curva ROC del modelo. La diagonal representa el clasificador aleatorio (AUC = 0,5).

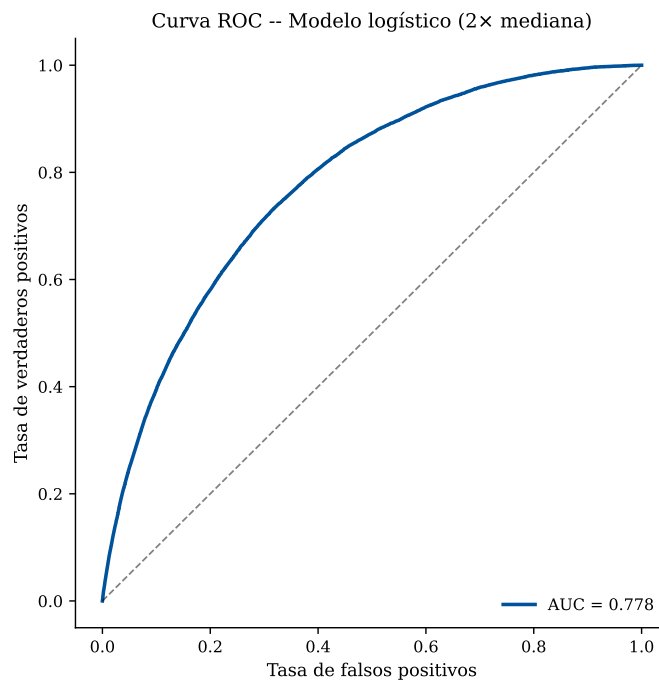


Figura 4.3: Curva ROC del modelo logístico ponderado. AUC = 0,778. La diagonal representa el clasificador aleatorio. Fuente: EPF microdatos INE, 2016–2024. Elaboración propia.

La tabla 4.4 presenta los factores de inflación de la varianza (VIF) de los predictores del modelo. Todos los valores se sitúan por debajo de 10, lo que indica que no hay multicolinealidad entre predictores.

Cuadro 4.4: Factor de inflación de la varianza (VIF)

Variable	VIF
Quintil 1 gasto	7,31
Quintil 2 gasto	3,64
Ocupado por cuenta ajena temp.	2,93
Quintil 3 gasto	2,44
Zona poco poblada	2,30
Edif. 10+ viviendas	2,22
Pareja con 2 hijos	2,17
Propiedad con hipoteca	1,99
Quintil 4 gasto	1,96
Persona sola <65	1,94
Edif. <10 viviendas	1,80
Zona intermedia	1,79
Antes de 1980	1,71
Persona sola 65+	1,59
Pareja con 3+ hijos	1,55
Pareja sin hijos ambos 65+	1,50
Vivienda unifamiliar	1,46
Alquiler a precio de mercado	1,45
Piso	1,45
Parado (busca primer empleo)	1,43
Adulto con hijos	1,43
Otro tipo de hogar (3+ adultos)	1,42
Persona sola (otro)	1,36
Jubilado	1,27
Pareja con 1 hijo	1,15
Vivienda independiente	1,13
Cesion gratuita	1,08
Otra forma de tenencia	1,08
Otro tipo de hogar (2 adultos)	1,08
Ocupado por cuenta ajena indef.	1,06
Alquiler inferior al mercado	1,05
Incapacitado permanente	1,05
Pareja sin hijos	1,04
Otra situacion	1,04
Otro tipo de edificio	1,01
Parado (ha trabajado)	1,01

Nota: CCAA y year dummies excluidos.

4.3.4. Análisis de sensibilidad

La definición del umbral que separa la carga energética excesiva de la sostenible es una convención aplicada en este estudio. Para verificar que los resultados no dependen de esa elección, se reestima el modelo con tres umbrales alternativos: 10 % fijo (transposición del umbral de Hills/Boardman al eje gasto), $2 \times$ mediana (umbral principal, 7,8 %) y percentil 75 muestral ($\approx 5,9\%$).

En esta tabla se reporta el OR y no el AME, a diferencia del cuadro principal del modelo (4.3). La razón es que los tres umbrales que se comparan generan situaciones con distinta frecuencia. Por eso, el AME en puntos porcentuales cambia de forma casi automática según lo frecuente o poco frecuente que sea el evento en cada caso. En cambio, el OR, al ser una medida relativa, permite comparar mejor los resultados entre columnas sin que esa diferencia de frecuencia distorsione tanto la interpretación.

Los resultados de la tabla 4.5 muestran que la jerarquía de los predictores se mantiene estable en las tres especificaciones: los quintiles bajos de gasto total siguen siendo el predictor dominante, el alquiler inferior al mercado ocupa el segundo lugar, y los efectos de densidad, tipo de edificio y composición del hogar conservan sus signos y magnitudes relativas. Esta estabilidad indica que las conclusiones del modelo no son fruto de la convención adoptada.

Los valores numéricos varían de forma previsible: cuando el evento de interés se define como más restrictivo (la cola por encima del umbral del 10 % es más pequeña que la del 2M), los *odds ratios* de los predictores que lo discriminan tienden a crecer. La AUC oscila entre valores comparables en los tres umbrales, lo que confirma que la capacidad de predicción del modelo es robusta.

4.3.5. Perfil de hogar vulnerable

Los resultados del modelo permiten trazar un perfil del hogar con mayor probabilidad de presentar una carga energética excesiva. Se trata de un hogar con bajo nivel de gasto total (quintiles inferiores), en régimen de alquiler inferior al mercado, residente en un edificio pequeño o en un municipio de baja densidad, cuyo sustentador principal está sin empleo

Cuadro 4.5: Análisis de sensibilidad del modelo: OR por umbral de carga energética (errores estándar cluster-robust por CCAA× año)

Variable	10%	2× mediana	P75
Quintil 1 gasto	17,340***	13,230***	11,900***
Quintil 2 gasto	7,907***	6,463***	5,832***
Quintil 3 gasto	4,705***	3,877***	3,633***
Alquiler inferior al mercado	3,162***	2,673***	2,231***
Quintil 4 gasto	2,711***	2,327***	2,229***
Edif. <10 viviendas	2,242***	2,076***	1,879***
Alquiler a precio de mercado	2,056***	1,849***	1,685***
Zona poco poblada	2,061***	1,829***	1,689***
Persona sola <65	2,005***	1,672***	1,427***
Parado (busca primer empleo)	1,592***	1,582***	1,465***
Edif. 10+ viviendas	1,732***	1,581***	1,480***
Jubilado	1,556***	1,539***	1,454***
Zona intermedia	1,574***	1,487***	1,413***
Otra situacion	1,410**	1,378***	1,353***
Persona sola (otro)	1,306***	1,329***	1,241***
Pareja con 1 hijo	0,766**	0,715***	0,694***
Incapacitado permanente	1,366***	1,313***	1,163***
Vivienda unifamiliar	1,404***	1,305***	1,232***
Pareja con 2 hijos	1,450***	1,297***	1,231***
Pareja sin hijos ambos 65+	0,796***	0,755***	0,709***
Ocupado por cuenta ajena temp.	1,273***	1,280***	1,232***
Otro tipo de hogar (2 adultos)	0,769*	0,766***	0,774***
Otro tipo de hogar (3+ adultos)	1,271***	1,236***	1,103***
Pareja con 3+ hijos	0,923	0,825***	0,805***
Ocupado por cuenta ajena indef.	1,132*	1,179**	1,132**
Adulto con hijos	0,883*	0,838***	0,810***
Persona sola 65+	1,275***	1,157***	1,068*
Otra forma de tenencia	1,177**	1,129	1,001
Pareja sin hijos	1,230	1,111	1,073
Cesion gratuita	1,318***	1,108	1,122***
Parado (ha trabajado)	1,092	1,092	0,790
Propiedad con hipoteca	1,057	1,062**	1,090***
Piso	0,977	1,032	1,050
Otro tipo de edificio	1,433	1,021	1,016
Antes de 1980	0,989	0,984	0,995
Vivienda independiente	0,970	1,004	1,018

Nota: errores estándar cluster-robust por CCAA× año; pesos muestrales normalizados. Se reporta el OR (no el AME) porque permite comparar directamente columnas entre umbrales con prevalencias distintas; el AME varía mecánicamente con la probabilidad base del evento. Dummies de CCAA y año incluidas en el modelo pero excluidas del listado. *** p<0,01; ** p<0,05; * p<0,1.

o jubilado, y que suele ser unipersonal.

El peso del gasto total sobre la probabilidad de vulnerabilidad es el hallazgo más sólido: los hogares del quintil más bajo tienen una probabilidad 36,5 p.p. mayor que los del quintil más alto (OR = 13,230). El régimen de tenencia aporta un efecto sustancial e independiente del nivel de gasto; el alquiler inferior al mercado eleva la probabilidad en 12,2 p.p. (OR = 2,673). Los resultados se mantienen además al cambiar el umbral de definición, lo que descarta que los hallazgos dependan de la convención escogida.

El capítulo 5 traduce estos hallazgos en implicaciones para la política pública y para el diseño de productos en el sector asegurador.

Capítulo 5

Conclusiones

5.1. Principales hallazgos

El nivel de gasto total del hogar es el factor que más pesa en el modelo. Los hogares del primer quintil de gasto tienen una probabilidad 36,5 p.p. mayor de sufrir carga energética excesiva que los del quintil de referencia. Esa diferencia se reduce en el segundo quintil (23,7 p.p.) y en el tercero (16,1 p.p.). La idea de fondo es sencilla: cuando el presupuesto total del hogar es más bajo, una parte mayor se destina a gastos básicos como la energía, que además son difíciles de recortar.

El régimen de tenencia es el segundo factor más relevante. Los hogares en alquiler por debajo del precio de mercado tienen una probabilidad 12,2 p.p. mayor de carga energética excesiva que los de la categoría de referencia. Esto sugiere que estos hogares ya parten de una situación económica más frágil y, además, al no ser propietarios, tienen menos margen para hacer mejoras de eficiencia en la vivienda.

Las características físicas de la vivienda también influyen de forma bastante clara. Vivir en edificios con menos de diez viviendas añade 8,2 p.p. de probabilidad, y hacerlo en edificios de diez o más viviendas añade 4,7 p.p. Esto encaja con la idea de que los bloques pequeños suelen perder más calor por vivienda. Además, la baja densidad del municipio suma 6,4 p.p., lo que puede estar relacionado con una mayor dependencia de combustibles fuera de red en zonas rurales. En cambio, la antigüedad de la vivienda, medida con una

variable dummy para viviendas anteriores a 1980, apenas aporta 0,16 p.p. y no resulta estadísticamente significativa ($p = 0,62$). Es decir, una vez controlado el resto, no aparece un efecto diferencial claro.

También destaca el tipo de hogar. Las personas que viven solas y tienen menos de 65 años presentan una probabilidad 5,5 p.p. mayor que la categoría de referencia. Una posible explicación es que combinan un gasto total relativamente bajo con unos consumos energéticos mínimos que no se reparten entre varios miembros del hogar.

La situación laboral del sustentador principal también marca diferencias. Los hogares cuyo sustentador busca su primer empleo suman 4,9 p.p., y aquellos cuyo sustentador está jubilado 4,6 p.p. En ambos casos, el resultado parece coherente con una menor capacidad para absorber subidas en la factura energética.

Las diferencias territoriales también son relevantes. Ceuta y Melilla presentan una probabilidad 9,8 p.p. mayor que la comunidad autónoma de referencia, lo que las sitúa como los territorios con mayor riesgo relativo. Este resultado puede estar recogiendo tanto sus mayores costes energéticos como características propias del parque residencial.

En conjunto, el modelo alcanza un AUC de 0,778 y un pseudo- R^2 de McFadden de 0,145, lo que apunta a una capacidad predictiva moderada. Este nivel de ajuste es razonable en datos de encuesta, donde siempre hay factores no observados que limitan la capacidad explicativa del modelo. Además, el análisis de sensibilidad con umbrales alternativos (10 % fijo y percentil 75) confirma que los principales resultados se mantienen: los signos y el peso relativo de los predictores cambian poco, lo que refuerza la solidez del diagnóstico.

5.2. Implicaciones para el sector asegurador

Este Trabajo Fin de Grado forma parte de un proyecto más amplio sobre sanidad financiera de los hogares españoles, en el que otros trabajos analizan dimensiones como la renta, el endeudamiento o la liquidez. La aportación de este trabajo se centra en la palanca energética y permite identificar perfiles de hogar cuya vulnerabilidad puede observarse, al menos en parte, en la relación entre aseguradora y asegurado. Esto es importante porque

varias de las variables que el modelo señala como relevantes (régimen de tenencia, tipo de edificación, composición del hogar o zona geográfica) coinciden con información que ya suele estar presente en pólizas o procesos de suscripción.

A partir de los resultados del modelo se plantean tres posibles líneas de aplicación. Su utilidad real, en todo caso, tendría que validarse con datos internos de una aseguradora, como historial de siniestralidad, tasas de renovación o elasticidad-precio de la demanda, algo que queda fuera del alcance de este trabajo.

5.2.1. Segmentación de primas de hogar

El gradiente de riesgo por quintil de gasto total es la señal más clara del modelo. Un hogar del primer quintil presenta una probabilidad 36,5 p.p. superior a la de un hogar del quintil de referencia; en el segundo quintil esa diferencia baja a 23,7 p.p. y en el tercero a 16,1 p.p. Esto sugiere que la proporción de gasto energético sobre el gasto total podría servir como variable complementaria en la tarificación de seguros de hogar. Dos hogares con características similares en la vivienda pueden tener perfiles de riesgo energético muy distintos si su estructura de gasto también lo es. Esa diferencia no siempre queda bien recogida en los modelos actuariales tradicionales, que suelen centrarse más en el inmueble que en la situación presupuestaria del hogar. La información necesaria para aproximar esta variable podría integrarse en los métodos de cálculo de las pólizas.

5.2.2. Producto contra picos de gasto energético

Los hogares en alquiler por debajo del mercado combinan una probabilidad 12,2 p.p. mayor de carga excesiva con poca capacidad de reacción ante una subida brusca de la factura. Cuando el margen presupuestario es pequeño, ese shock puede traducirse en impago o en renuncia a niveles adecuados de confort térmico.

A partir de esto, una posible aplicación sería un seguro paramétrico, similar en lógica a los seguros agrarios por índice climático. El producto se activaría cuando el coste de la energía superase un umbral predefinido, por ejemplo una determinada desviación sobre la media móvil del IPC energético, y pagaría una cuantía fija a partir de este. La idea

económica es sencilla: ofrecer un colchón frente a shocks de factura a hogares que, por su situación presupuestaria, no pueden generarlo por sí solos.

5.2.3. Financiación de rehabilitación energética

Los edificios con menos de diez viviendas (8,2 p.p. adicionales) y la baja densidad del municipio (6,4 p.p.) apuntan a que la eficiencia térmica de la vivienda influye de forma relevante en la carga energética. A esto se suma el caso del alquiler por debajo de mercado (12,2 p.p.), donde el inquilino suele tener menos incentivos y menos capacidad real para invertir en mejoras del inmueble.

Un posible producto que ataca estas palancas sería un préstamo verde vinculado al seguro de hogar, con una bonificación en la prima si el asegurado realiza una rehabilitación elegible. La lógica es que una mejora térmica reduce de forma estructural el gasto energético y, por tanto, también la probabilidad de tensión financiera asociada a subidas de la factura. Para la aseguradora, además, esto podría favorecer la retención del cliente durante la vida del préstamo. Esta línea parece más viable en vivienda en propiedad, aunque a futuro podrían pensarse soluciones específicas también para el mercado del alquiler.

5.3. Limitaciones y líneas futuras

El análisis tiene varias limitaciones que conviene tener en cuenta al interpretar los resultados.

En primer lugar, la EPF es una encuesta transversal repetida (*repeated cross-section*), no un panel. Por tanto, el modelo estima la probabilidad de que un hogar esté en situación de carga energética excesiva en un momento dado, pero no permite estudiar bien cómo entra o sale de esa situación con el tiempo. Para analizar esas transiciones harían falta datos longitudinales, como los de la Encuesta de Condiciones de Vida (ECV) o el EU-SILC.

En segundo lugar, la variable dependiente se construye como gasto energético sobre gasto total, y no sobre renta. Esto se debe a que la EPF, al ser una encuesta de presupuestos, no recoge directamente la renta del hogar en sus ficheros de uso público. La medida utilizada

encaja bien con una lectura presupuestaria del problema, pero se aparta de los umbrales clásicos de la literatura, que suelen basarse en la renta.

Además, el modelo logístico binario permite identificar asociaciones entre variables, pero no relaciones causales. Por ejemplo, que el régimen de tenencia aparezca relacionado con una mayor carga energética no significa, por sí solo, que la cause. Para responder a ese tipo de preguntas harían falta diseños causales más exigentes, que quedan fuera del alcance de este trabajo.

En cuanto a líneas futuras, una de las más claras sería realizar un análisis longitudinal con el panel rotante de la ECV, lo que permitiría estudiar mejor las transiciones hacia situaciones de vulnerabilidad energética. También sería útil combinar la EPF con datos de consumo físico, en kWh, procedentes del IDAE o de distribuidoras, para separar mejor qué parte de la carga se debe a mayor consumo y qué parte a un mayor precio. Por otro lado, la dimensión territorial, que aquí solo se recoge mediante efectos fijos, podría analizarse con más detalle mediante modelos más complejos si se contara con suficientes datos por territorio. Por último, las aplicaciones aseguradoras planteadas en este trabajo necesitarían una validación posterior con datos internos de una aseguradora, algo que no es posible abordar con fuentes públicas.

Bibliografía

- Boardman, B. (1991). *Fuel Poverty: From Cold Homes to Affordable Warmth*. Belhaven Press.
- Boardman, B. (2010). *Fixing Fuel Poverty: Challenges and Solutions*. Earthscan.
- Bouzarovski, S. (2018). *Energy Poverty: (Dis)Assembling Europe's Infrastructural Divide*. Palgrave Macmillan. <https://doi.org/10.1007/978-3-319-69299-9>
- Bouzarovski, S., & Petrova, S. (2015). A global perspective on domestic energy deprivation: Overcoming the energy poverty–fuel poverty binary. *Energy Research & Social Science*, 10, 31-40. <https://doi.org/10.1016/j.erss.2015.06.007>
- Costa-Campi, M. T., Duch-Brown, N., & Trujillo-Baute, E. (2019). Energy Poverty in Spain: Trends, Determinants and Policy Implications. *Energy Policy*, 128, 959-969. <https://doi.org/10.1016/j.enpol.2018.12.039>
- Deaton, A., & Muellbauer, J. (1980). *Economics and Consumer Behaviour*. Cambridge University Press.
- Dubois, U., & Meier, H. (2016). Energy Affordability and Energy Poverty in Europe: Re-assessing the Relationships. *Energy Research & Social Science*, 18, 21-35. <https://doi.org/10.1016/j.erss.2016.04.006>
- Energy Poverty Advisory Hub. (2023). *Indicators and Data on Energy Poverty in Europe* (inf. téc.). European Commission Energy Poverty Advisory Hub (EPAH). Brussels. <https://energy-poverty.ec.europa.eu/>
- Eurostat. (2026). Inability to Keep Home Adequately Warm, by Household Type (ILC_MDES01) [Datos 2004–2025, geo=ES, actualización 27/02/2026. DOI: 10.2908/ILC_MDES01]. https://ec.europa.eu/eurostat/databrowser/view/ILC_MDES01/

- Healy, J. D. (2004). *Housing, Fuel Poverty and Health: A Pan-European Analysis*. Ashgate.
- Hernández, D. (2016). Understanding 'energy insecurity' and why it matters to health. *Social Science & Medicine*, 167, 1-10. <https://doi.org/10.1016/j.socscimed.2016.08.029>
- Hills, J. (2012). *Getting the Measure of Fuel Poverty: Final Report of the Fuel Poverty Review* (CASE report 72). Department of Energy y Climate Change. London. <https://sticerd.lse.ac.uk/dps/case/cr/CASEREport72.pdf>
- Legendre, B., & Ricci, O. (2015). Measuring fuel poverty in France: Which households are the most fuel vulnerable? *Energy Economics*, 49, 620-628. <https://doi.org/10.1016/j.eneco.2015.01.022>
- Liddell, C., & Morris, C. (2010). Fuel Poverty and Human Health: A Review of Recent Evidence. *Energy Policy*, 38(6), 2987-2997. <https://doi.org/10.1016/j.enpol.2010.01.037>
- Maxim, A., Mihai, C., Apostoaie, C.-M., Popescu, C., Istrate, C., & Bostan, I. (2016). Implications and Measurement of Energy Poverty Across the European Union. *Sustainability*, 8(5), 483. <https://doi.org/10.3390/su8050483>
- Middlemiss, L. (2017). A Critical Analysis of the New Politics of Fuel Poverty in England. *Critical Social Policy*, 37(3), 425-443. <https://doi.org/10.1177/0261018316674851>
- Middlemiss, L., & Gillard, R. (2015). Fuel poverty from the bottom-up: Characterising household energy vulnerability through the lived experience of the fuel poor. *Energy Research & Social Science*, 6, 146-154. <https://doi.org/10.1016/j.erss.2015.02.001>
- Ministerio para la Transición Ecológica y el Reto Demográfico. (2019). *Estrategia Nacional contra la Pobreza Energética 2019–2024* (inf. téc.) (Estrategia Nacional contra la Pobreza Energética 2019–2024. MITECO.). Gobierno de España. Madrid. <https://www.miteco.gob.es/es/energia/eficiencia/>
- Moore, R. (2012). Definitions of Fuel Poverty: Implications for Policy. *Energy Policy*, 49, 19-26. <https://doi.org/10.1016/j.enpol.2012.01.057>
- Okushima, S. (2016). Measuring Energy Poverty in Japan, 2004–2013. *Energy Policy*, 98, 557-564. <https://doi.org/10.1016/j.enpol.2016.09.005>

- Pelz, S., Pachauri, S., & Groh, S. (2018). A Critical Review of Modern Approaches for Multidimensional Energy Poverty Measurement. *Wiley Interdisciplinary Reviews: Energy and Environment*, 7(6), e304. <https://doi.org/10.1002/wene.304>
- Romero, J. C., Linares, P., & López-Otero, X. (2014). *Pobreza energética en España: Análisis económico y propuestas de actuación* (inf. téc. N.º WP 08/2014). Economics for Energy. Vigo.
- Romero, J. C., Linares, P., & López-Otero, X. (2018). The policy implications of energy poverty indicators. *Energy Policy*, 115, 98-108. <https://doi.org/10.1016/j.enpol.2017.12.054>
- Sánchez-Guevara Sánchez, C., Sánchez de la Flor, F. J., Sanz Fernández, A., Sanz Fernández, T., & Fernández Ans, P. (2019). Energy Poverty in Madrid: Adaptive Comfort Approach. *Building and Environment*, 157, 54-63. <https://doi.org/10.1016/j.buildenv.2019.04.019>
- Thomson, H., Bouzarovski, S., & Snell, C. (2017). Rethinking the measurement of energy poverty in Europe: A critical analysis of indicators and data. *Indoor and Built Environment*, 26(7), 879-901. <https://doi.org/10.1177/1420326X17699260>
- Thomson, H., & Snell, C. (2013). Quantifying the Prevalence of Fuel Poverty Across the European Union. *Energy Policy*, 52, 563-572. <https://doi.org/10.1016/j.enpol.2012.10.009>
- Tirado Herrero, S., Jiménez Meneses, L., López Fernández, J. L., & Irigoyen Hidalgo, V. M. (2018). *Pobreza Energética en España 2018: Hacia un sistema de indicadores y una estrategia de actuación estatales* (inf. téc.). Asociación de Ciencias Ambientales (ACA). Madrid.
- Tirado Herrero, S., & Ürge-Vorsatz, D. (2012). Trapped in the Heat: A Post-Communist Building Stock as a Source of Energy Poverty. *Energy Policy*, 49, 2-11.

Apéndice A

Declaración de Uso de Herramientas de Inteligencia Artificial Generativa en Trabajos Fin de Grado

ADVERTENCIA: Desde la Universidad consideramos que ChatGPT u otras herramientas similares son herramientas muy útiles en la vida académica, aunque su uso queda siempre bajo la responsabilidad del alumno, puesto que las respuestas que proporciona pueden no ser veraces. En este sentido, NO está permitido su uso en la elaboración del Trabajo fin de Grado para generar código porque estas herramientas no son fiables en esa tarea. Aunque el código funcione, no hay garantías de que metodológicamente sea correcto, y es altamente probable que no lo sea.

Por la presente, yo, Gonzalo Alonso Lidón, estudiante de Máster en Tecnologías de Telecomunicaciones y Business Analytics de la Universidad Pontificia Comillas al presentar mi Trabajo Fin de Grado titulado “Determinantes de la carga energética excesiva en los hogares españoles a partir de microdatos de la EPF 2016–2024”, declaro que he utilizado la herramienta de Inteligencia Artificial Generativa ChatGPT u otras similares de IAG de código sólo en el contexto de las actividades descritas a continuación:

1. **Brainstorming de ideas de investigación:** Utilizado para idear y esbozar posibles áreas de investigación.

2. **Crítico:** Para encontrar contra-argumentos a una tesis específica que pretendo defender.
3. **Referencias:** Usado conjuntamente con otras herramientas, como Science, para identificar referencias preliminares que luego he contrastado y validado.
4. **Metodólogo:** Para descubrir métodos aplicables a problemas específicos de investigación.
5. **Interpretador de código:** Para realizar análisis de datos preliminares.
6. **Estudios multidisciplinares:** Para comprender perspectivas de otras comunidades sobre temas de naturaleza multidisciplinar.
7. **Constructor de plantillas:** Para diseñar formatos específicos para secciones del trabajo.
8. **Corrector de estilo literario y de lenguaje:** Para mejorar la calidad lingüística y estilística del texto.
9. **Generador previo de diagramas de flujo y contenido:** Para esbozar diagramas iniciales.
10. **Sintetizador y divulgador de libros complicados:** Para resumir y comprender literatura compleja.
11. **Generador de problemas de ejemplo:** Para ilustrar conceptos y técnicas.
12. **Revisor:** Para recibir sugerencias sobre cómo mejorar y perfeccionar el trabajo con diferentes niveles de exigencia.
13. **Traductor:** Para traducir textos de un lenguaje a otro.

Afirmo que toda la información y contenido presentados en este trabajo son producto de mi investigación y esfuerzo individual, excepto donde se ha indicado lo contrario y se han dado los créditos correspondientes (he incluido las referencias adecuadas en el TFG y he explicitado para que se ha usado ChatGPT u otras herramientas similares). Soy consciente de las implicaciones académicas y éticas de presentar un trabajo no original y acepto las consecuencias de cualquier violación a esta declaración.

Fecha: 21.04.2026

Firma: Gonzalo Alonso Lidón