

Tipo de documento: Tesis de Maestría



Departamento de Economía. Maestría en Econometría

Estimando el Impacto de la Sustitución del Gas Licuado del Petróleo por el Gas Natural en Lima Metropolitana

Autoría: Rivas Castillo, Juan Manuel

Año: 2024

¿Cómo citar este trabajo?

Rivas Castillo, J. (2024) "Estimando el Impacto de la Sustitución del Gas Licuado del Petróleo por el Gas Natural en Lima Metropolitana". [Tesis de Maestría. Universidad Torcuato Di Tella]. Repositorio Digital Universidad Torcuato Di Tella

<https://repositorio.utdt.edu/handle/20.500.13098/13207>

El presente documento se encuentra alojado en el Repositorio Digital de la Universidad Torcuato Di Tella bajo una licencia Creative Commons Atribución-No Comercial-Compartir Igual 4.0 Argentina (CC BY-NC-SA 4.0 AR)
Dirección: <https://repositorio.utdt.edu>



UNIVERSIDAD
TORCUATO DI TELLA

Estimando el Impacto de la Sustitución del Gas Licuado del Petróleo por el Gas Natural en Lima Metropolitana

Autor: Juan Manuel Rivas Castillo

Legajo: 21A1288

Asesor: Martín Gonzalez Rozada

Departamento de Economía

Maestría en Econometría

30 de mayo de 2024

Índice

1. INTRODUCCION	3
2. MARCO TEÓRICO Y EMPÍRICO	5
2.1. Modelando la Elección del Combustible para Cocinar	5
2.2. Modelando el Gasto en el Combustible para Cocinar	7
2.3. Proceso de Difusión del GN	9
2.4. Aproximación a la estimación del Impacto en el Gasto de los Hogares de Lima Metropolitana	11
3. ANÁLISIS Y ESTIMACIONES	14
3.1. Análisis de la Información	14
3.2. Estimación del Proceso de Difusión del GN en Lima Metropolitana	17
3.3. Estimación del Impacto	20
3.3.1. Diferencia de Medias	20
3.3.2. Estimación Mediante el Método de Horváth y Maldoom	24
3.3.3. Estimación Mediante Propensity Score Mat- ching	25
4. Resultados Finales	28
5. Conclusiones	28
Referencias	29
6. ANEXOS	32

Resumen

Entre los años 2007 y 2023, el total de hogares que empleaban Gas Natural como insumo para coccion de alimentos en Lima Metropolitana, se multiplicó por 78; mientras que para el mismo periodo de tiempo, el total de hogares que emplean el balón de Gas Licuado de Petróleo se multiplicó solamente por 1.1. Si bien aún la mayoría de hogares usan como combustible principal para cocinar el balón de Gas Licuado de Petróleo, el crecimiento en el uso del Gas Natural resulta en un evento relevante que esta modificando el presupuesto del hogar destinadao para la cocción de alimentos; por lo que en esta investigación se estimó la cuantía del impacto en el hogar de este proceso de sustitución de combustibles.

Para estimar el impacto se empleó información soecioeconómica de 17 cortes trasnversales, a partir de la Encuesta Nacional de Hogares para el periodo 2007 al 2023. Con esta información se estimaron 7 especificaciones del impacto monetario para cada uno de los años, obteniéndose resultados similares con el empleo de técnicas de *Propensity Score Matching*. Además, mediante información real de la evolución de la adopción del Gas Natural en Lima Metropolitana, se estimó el proceso de adopción de este combustible; por lo que se asociaron los resultados del impacto con la fase en la cual se encuentra el nivel de adopción en Lima Metropolitana.

Los resultados de las estimaciones efectuadas permiten concluir que los hogares de Lima Metropolitana en la actualidad se encuentran en la penúltima fase de adopción denominada *mayoría tardia*, con lo cual obtuvieron un ahorro acumulado de S/ 24 254 652 (US\$ 6 565 401), producto del empleo del Gas Natural como combustible para cocinar en lugar del balón de Gas Licuado de Petróleo.

Palabras Clave : Combustibles, Cocinar, Hogares, Efecto, Sustitución, Ingreso, Elección, Evaluación de Impacto, Modelo de Difusión Tecnológica

1. INTRODUCCION

Según la Encuesta Nacional de Hogares (ENAHO), elaborada por el Instituto Nacional de Estadística e Informática (INEI)¹, al 2007 habían un total de 2 millones de hogares en Lima Metropolitana², de los cuales el 0.8% empleaba como combustible principal para cocinar sus alimentos el Gas Natural (GN); mientras que el 84.6%, balones de Gas Licuado de Petróleo (GLP)³.

Para el 2023, el total de hogares en Lima Metropolitana se duplicó; mientras que el 53% emplean GLP y el 32% GN. Respecto de su evolución comparando el 2007, el total de hogares que emplean GLP se multiplicó por un factor de 1.1; y los que emplean GN para cocinar por un factor de 78. Esta evolución revela evidencia preliminar de que no solamente los nuevos hogares adoptan el GN como insumo para cocinar sus alimentos, sino que también viene ocurriendo un proceso de sustitución del balón de GLP por el GN.⁴

Este crecimiento en el proceso de adopción del GN motiva a la presente investigación; pues la pregunta principal a responder es: ¿Cuál es el impacto monetario en los hogares de Lima Metropolitana que emplean el GN como combustible principal para la cocción de alimentos?. Como un punto de partida, y en la búsqueda de una respuesta, se puede afirmar que el impacto se asocia con la fase de adopción del GN. Al respecto, (Bass, 1969) permite modelar el proceso de difusión por el cual pasa un nuevo producto, este proceso de manera empírica suele formar una curva en forma de S, que es aproximada mediante funciones de distribución acumuladas; de manera general, podemos afirmar que las etapas son: introducción, expansión y saturación del producto.

Un detalle no menor es que el crecimiento del GN en Lima Metropolitana fue impulsado a partir de programas de gobierno que financian el acceso a redes y puntos de conexión. Al respecto, el Programa Bonogas, que inició el 2016, financia el acceso a aquellos hogares que pertenecen a estratos socioeconómicos bajos⁵.

¹Los módulos de la encuesta se encuentran disponibles en el siguiente link <https://proyectos.inei.gob.pe/microdatos/>

²Lima Metropolitana se encuentra conformado por la provincia de Lima, que cuenta con 43 distritos, y la de la provincia del Callao, con 7 distritos.

³Cabe notar que la información reportada por la ENAHO se encuentra subestimada pues señala que al 2022 cerca de un millón de hogares en Lima Metropolitana cuentan con GN; no obstante, según información de la Dirección General de Hidrocarburos del Ministerio de Energía y Minas, para fines del 2022, cerca de un millón y medio de hogares contaba con acceso al GN

⁴El proceso de sustitución del balón de GLP por el GN no solamente requiere de la disponibilidad de una cocina, sino también el despliegue de redes y puntos de conexión.

⁵Este programa otorga financiamiento a hogares con bajos recursos para que se construyan conexiones de GN al interior de sus viviendas. El requisito principal para acceder al Bonogas fue que la vivienda se debe encontrar ubicada en los planos estratificados del INEI con categoría socioeconómica de estrato medio, medio-bajo y bajo. El hogar al que se le brindó este benefi-

Para la estimación del impacto se requiere comprender que existe un proceso en dos etapas; en la primera de ellas, los hogares eligen el combustible que emplearan para cocinar; mientras que en la segunda etapa, el gasto en el empleo del combustible dependerá de las características con las que cuente el hogar. Respecto a la primera etapa, la literatura especializada señala que los hogares valoran factores como el precio, el ingreso y las costumbres al momento de elegir el combustible para cocinar. De otro lado, el gasto depende de factores como el tamaño del hogar, el nivel educativo, la edad del jefe del hogar, entre otras.

Para el Perú no se identifica un antecedente de investigación que estime el impacto en el gasto de los hogares de Lima Metropolitana debido al empleo del GN comparado con el balón del GLP. No obstante, en la página web de la empresa Calidda, distribuidora del GN en Lima Metropolitana, se cuenta con un simulador que calcula el ahorro que obtendría el hogar si en lugar de emplear un balón de GLP emplea GN, el resultado observado a la fecha de consulta es de S/ 8.62⁶. Valor que se calcula empleando información comparativa del valor de un balón de GLP de 10 kg de la página web Facilito.⁷

No obstante, el cálculo realizado por la empresa es directo sin considerar las características socioeconómicas de los hogares, los patrones de gasto y el proceso de difusión del GN en Lima Metropolitana. Por ello, para poder estimar el impacto de adoptar el GN, en esta investigación se emplean distintos módulos de la ENAHO para el periodo 2007 al 2023. Esto con el objeto de poder determinar el impacto diferenciado, en función de las fases del proceso de adopción del GN en Lima Metropolitana.

De acuerdo con lo descrito, esta investigación resulta inédita; por lo que se espera que contribuya, no solamente con la discusión académica de como modelar el impacto en el gasto de los hogares producto del consumo de GN; sino que además ayude, con evidencia empírica, a los hacedores de política para implementar medidas de acceso al GN en todo el país.

cio debe devolver el financiamiento, de manera diferenciada y dependiendo al estrato al cuál pertenecen, hasta en 10 años, sin intereses ni cuota inicial a través de los recibos mensuales de consumo de GN. De acuerdo con ello, los de estrato bajo no devuelven el financiamiento; mientras que los del medio bajo devuelven el 25 % y los del medio, el 50 %.

⁶Consulta realizada en mayo del 2024 en el siguiente link <https://www.calidda.com.pe/mi-hogar/simulador-de-ahorro>

⁷Página disponible en el siguiente link: <https://www.facilito.gob.pe/facilito/pages/facilito/buscadorEESS.jsp>

2. MARCO TEÓRICO Y EMPÍRICO

2.1. Modelando la Elección del Combustible para Cocinar

Para modelar la elección del combustible para cocinar en Lima Metropolitana, una referencia que no se puede dejar de lado es la de (Rivas, 2020), quien empleando modelos logit multinomiales sucesivos estimó la demanda de los hogares de Lima Metropolitana por combustibles para cocinar. La especificación empleada por este autor fue la siguiente:

$$Pr(Y = Combustible_i) = \frac{e^{V_{ci}}}{\sum_{i=1}^m e^{V_{ci}}} \quad (1)$$

De la expresión anterior se puede observar que la demanda depende de una función de utilidad indirecta (V_i) de los combustibles estudiados. Función que se construye en función de un conjunto de variables observables que se asocian con la elección de cada uno de los combustibles empleados para cocinar.

De la revisión de literatura, el autor identifica que existen variables que establecen una relación de tipo estructural con la probabilidad de elección del combustible para cocinar, estas variables son el precio del combustible y el ingreso del hogar. El autor advierte que de ser omitidas estas variables la estimación generarían problemas de endogeneidad. Dado que en la ENAHO no se cuenta con el precio del combustible empleado para cocinar, el autor lo aproxima mediante el gasto mensual del hogar.

Además de las variables identificadas, el autor incluyó algunas variables de control adicionales, como por ejemplo: miembros del hogar, edad del jefe del hogar, sexo del jefe del hogar, que el hogar tenga hijos pequeños, entre otras. Entre el resultado más importante que identificó el autor es que un incremento del ingreso de hogares que cuentan con GN reduce la probabilidad el empleo del GLP.

En esa misma línea, (Osinermin, 2018) a partir de la metodología implementada por (Berry, Levinsohn, y Pakes, 1993) estimaron que existe sustitución de un solo lado; es decir, un incremento en el precio del balón de GLP en Lima Metropolitana genera un efecto de sustitución con el GN, pero lo contrario no ocurriría.

Una variable discutida con mucha regularidad para analizar la demanda de combustibles para cocinar; además de la sustitución de estos, es el ingreso del hogar. Al respecto (Hosier y Dowd, 1987) plantearon la hipótesis de la energía líder, hipótesis que establece una relación directa entre el ingreso y el combustible que los hogares seleccionan; por lo que un incremento en el ingreso impulsa la transición

hacia fuentes de energía más limpias y eficientes. De acuerdo con ello, hogares con bajos niveles de ingreso, tienden a usar energías baratas y disponibles localmente.

La hipótesis de la energía líder se encuentra relacionada con el modelo de la Escalera Energética, el cual describe el comportamiento de los hogares ante la decisión de consumir uno o varios tipos de combustibles según los niveles de renta que perciban (Van der Kroon, Brouwer, y Van Beukering, 2013). De acuerdo con ello, se ordenan los combustibles dependiendo de varias características; entre ellas, las propiedades físicas del combustible, la practicidad de su uso, la rapidez de cocción y la eficiencia (Hiemstra-Van der Horst y Hovorka, 2008).

Sin embargo, (Matera, Saatkamp, y Kammen, 2000) señala que la sustitución de combustibles dista de ser un proceso suave impulsado por el incremento de los ingresos; en lugar de ello, es el resultado de la interrelación de condiciones macroeconómicas estructurales, la cultura local y las circunstancias económicas de los hogares. También añaden que la adopción y uso de GLP se caracteriza por ser un proceso en dos etapas. En la primera, los hogares deben superar la barrera de no contar con una cocina para el empleo de GLP; mientras que en la segunda etapa, el consumo dependerá de las condiciones demográficas y de los requerimientos de energía vinculadas con los alimentos que se preparan de manera local.

En esa misma línea (Mekonnen y Köhlin, 2009) empleando un modelo de panel de datos de elección discreta con efectos aleatorios, encuentran que los hogares del África Subsahariana no cambian a combustibles más limpios cuando sus ingresos aumentan; en lugar de ello, tienden a incrementar la cantidad de opciones de combustibles para cocinar (*fuel stacking*).

Otra variable incluida en las investigaciones revisadas es el tamaño del hogar, la cual se calcula en función del número de miembros del hogar. Esta variable plantea un doble efecto; pues una mayor cantidad de miembros del hogar impulsa el empleo de combustibles tradicionales; y al mismo tiempo, representa un costo de adopción per cápita menor de un combustible más eficiente. Sin embargo, el efecto del factor tamaño del hogar no es uniforme en las investigaciones revisadas. (Hosier y Dowd, 1987) encuentran que resulta más probable que hogares grandes empleen el kerosene en lugar de la leña; pero menos probable que estos se cambien a electricidad. (Heltberg, 2004) señala que el tamaño del hogar no es una variable que afecte la probabilidad que una familia use leña en Guatemala; sin embargo, hogares pequeños son más probables que empleen GLP de manera exclusiva. En esa misma línea, (Ngui, Mutua, Osiolo, y Aligula, 2011) señalan que un incremento en el tamaño del hogar trae como consecuencia el cambio hacia fuentes de energía como: carbón, leña y GLP, ello con el objeto de compensar el incremento de la demanda de energía. Además, (Kroon, Brouwer, y Beukering, 2013) y (Rao y Reddy, 2007) afirman que hogares de un gran tamaño y con muchas mujeres, significan un bajo costo de oportunidad de recoger leña; por lo que ello conduce a que se almacene el combustible.

De otro lado, son pocos los estudios que emplean la variable precio como un factor determinante de la elección por parte del hogar del combustible para cocinar; sin embargo, se debe tomar en cuenta que el precio es empleado como una variable clave por los gobiernos para reorientar el consumo hacia determinados combustibles. (Barnes, Krutilla, y Hyde, 2010) señalan que los gobiernos pueden influenciar el empleo de ciertos combustibles por medio de dos canales: precios y accesibilidad. (Gupta y Köhlin, 2006) hacen notar que incrementar el precio de la leña y, a su vez, la disponibilidad de GLP es clave si lo que se busca es reducir la contaminación interna del hogar. (Mekonnen y Köhlin, 2009) encuentra que cuando el precio de la leña se incrementa los hogares de Etiopía cambian sus preferencias hacia el kerosene y electricidad.

De otro lado, en otra línea de investigación, y que tiene que ver con el concepto de la escalera energética, es el impacto ambiental negativo debido al uso de determinados combustibles. De acuerdo con esto último, para (Duffo, Greenstone, y Hanna, 2008), los combustibles para cocinar como la biomasa (leña y carbón) y los productos de petróleo (kerosene y GLP) tienen impactos ambientales negativos debido a la emisión de partículas en los hogares y el medio ambiente. (Geist y Lambin, 2002), señalan que el empleo de la leña como combustible para cocinar ha contribuido con el 28 % de la deforestación de los bosques tropicales ubicados en Asia, África y Latinoamérica.

En un estudio realizado para Colombia por (Alvis-Guzmán, Alvis-Estrada, y Orozco-Africano, 2008), quienes a partir de un análisis costo efectividad sobre el beneficio de la instalación de gas natural en reemplazo de las cocinas que funcionan con biomasa en una comunidad rural del Caribe Colombiano, estimaron que los costos evitados por carga de enfermedad atribuibles a la contaminación dentro de la vivienda, debido al uso del GN domiciliario, fueron menores en 30 % en relación a los hogares que utilizaban combustibles sólidos para cocción.

(Torres, Agreda, y Polo, 2010), manifiestan que, en las zonas rurales del Perú, el tipo de cocinas que utilizan para la cocción de sus alimentos son las llamadas cocinas tradicionales, siendo algunas de éstas las chimeneas, pero simples y de baja eficiencia, pues registran pérdidas normales entre el 80 % y 95 % de energía. Esta práctica que es esencial y vital para la población, es generadora de emisiones de gases de efecto invernadero contaminantes y que afectan la salud de los miembros del hogar.

2.2. Modelando el Gasto en el Combustible para Cocinar

El gasto que realiza un hogar en el combustible para cocinar se encuentra condicionado a la elección del combustible. Una de las aproximaciones para modelar el gasto de combustibles para cocinar se remonta a los trabajos de (G. S. Becker,

1965), quien argumentaba que los hogares pueden ser pensados como pequeñas fábricas que combinan inputs para la producción de bienes. En el caso de la producción de comida, un hogar típico emplea alguna fuente de energía que le sirve de insumo. De este modo, se podría plantear una función de producción para la cocción de alimentos del siguiente tipo:

$$f(\sigma k) \text{ con } f_k > 0 \text{ y } f_{kk} < 0 \quad (2)$$

Además, se considera el hecho de que los hogares no solo emplean sus ingresos para la cocción de comida preparada sino también para un conjunto de otros bienes a los cuales, por simplicidad, se les denomina el bien compuesto. De acuerdo con ello, la función de utilidad se representa de la siguiente manera:

$$\mu_i = \mu(z_i, x_i) = \mu(f(\sigma k), x_i) \quad (3)$$

La función Lagrangiana asociada es:

$$L = \mu(f(\sigma k), x_i) + \lambda(\beta - \alpha w^k K - x_i) \quad (4)$$

Tomando en cuenta una función de utilidad de tipo Cobb Douglas del tipo $\mu(z, x) = z^\theta x^{1-\theta}$ y $z = \sigma k$, se puede derivar el siguiente modelo empírico:

$$\ln(k_j) = \rho + \alpha_0 \ln(\beta_j) - \alpha_1 \ln(w_i) + \alpha_2 \ln(\theta_i) + \epsilon_i \quad (5)$$

De este modo el consumo de un determinado combustible para cocinar se encuentra en función directa del presupuesto del hogar y de las preferencias del hogar, pero en función inversa del costo del combustible.

A partir del modelo se puede notar que la disponibilidad de nuevas fuentes de energía para cocinar genera efectos en la restricción presupuestaria de los hogares, pues supone una modificación en los precios relativos de los combustibles. En el ámbito rural existen ciertos factores hacen que la sustitución sea lenta e incompleta; hecho que se explica por las costumbres propias de una determinada zona geográfica⁸. Aunque en el urbano, se emplea con mayor regularidad combustibles más eficientes (GLP, GN, electricidad). (Mekonnen y Köhlin, 2009) encuentran que los factores que determinan la elección del combustible para cocinar de los hogares ubicados en zonas urbanas de Etiopía son el gasto de los hogares, el precio y la educación.

⁸Un detalle identificado en la literatura es la lentitud, en ámbitos rurales, en la sustitución de combustibles para cocinar, (Ouedraogo, 2006)

Con el objeto de ampliar este enfoque se toma en cuenta lo señalado por (Soberon-Ferrer y Dardis, 1991), quienes identifican que el gasto en servicios de un hogar depende de variables asociadas a la producción del hogar, las cuales se miden a partir de los ingresos. De otro lado, también depende de variables que se relacionan con el ciclo de vida de la familia, las cuales se asocian con la edad del jefe de hogar, el número de hijos menores de 6 años y el total de personas en el hogar sin considerar a los niños menores de 6 años. La educación del jefe de hogar, que es una variable que se vincula con los gastos eficientes en el hogar. La ubicación del hogar, variable que captura diferencias en el gasto asociadas a vivir en una zona urbana o rural. Finalmente, el autor incluye la propiedad del hogar, pues se considera que los patrones de gasto cambian entre aquellos hogares que son propietarios de hogar respecto de los que no lo son.

En investigaciones internacionales, (Akpalu, Dasmani, y Aglobitse, 2011) plantean una especificación econométrica que relaciona el consumo de combustibles para cocinar, en kilojoules, con el precio, el ingreso y un conjunto de variables socioeconómicas. Sus resultados muestran elasticidades precio menores a uno para el caso el GLP, carbón y leña; mientras que para el kerosene, la elasticidad precio resultó mayor a uno. La explicación que brindan para la mayor sensibilidad del kerosene son los múltiples usos de este combustible.

A este conjunto de variables, identificadas en la literatura especializada, se le incluye la clasificación realizada por el INEI de pobreza monetaria. Variable que permite aproximar el subsidio que brinda el Estado Peruano a los hogares para acceder al GN según su nivel socioeconómico. De acuerdo con ello, es de esperarse que existan impactos diferenciados en el gasto respecto de si el hogar es pobre o no lo es.

2.3. Proceso de Difusión del GN

Los combustibles más eficientes generan efectos externos en la utilidad de los consumidores y, por ende, el consecuente incremento en la demanda por dicho bien. Esto contradice la teoría del consumidor, pues asume que los individuos toman sus decisiones de manera individual sin interactuar con otros.

Esta idea de interacción entre los consumidores se encuentra incorporada en el modelo de difusión de un nuevo producto propuesto por (Bass, 1969). Modelo que sirve para identificar las diferentes fases de como un nuevo producto es adoptado por una población específica; permitiendo el identificar la etapa en la cual se encuentre este.

A estos procesos de difusión, también denominados procesos acotados, se suelen emplear ajustes de tendencia no lineal (por ejemplo modelos Logit, Probit, Gompertz y cualquier otro que permita ajustar una curva en forma de S). No obstante, se requiere el conocer a priori el valor máximo de crecimiento del proceso de difusión (techo de referencia).

De manera general se puede afirmar que a partir de estos procesos de difusión se caracterizan las etapas de introducción, expansión y saturación de un nuevo producto. (Giovanis y Skiadas, 2007) afirman que si el proceso de difusión es estable, y sin influencias externas, es de esperar que el patrón de comportamiento siga un patrón de una curva en forma de S. De este modo, la curva en forma de S del proceso implica que en un inicio la adopción presenta una tendencia constante y luego es cada vez mayor hasta que alcanza el punto de inflexión; luego de ello, esta tendencia se reduce a medida que nos aproximamos al punto de saturación.

Como se indicó anteriormente, para estimar el modelo de difusión tecnológica es necesario conocer de antemano el valor del punto de saturación; es decir, el valor que ocasiona que la tasa de adopción del GN sea de cero. para calcular ello (Bass, 1969) parte del siguiente modelo de difusión de influencia mixta.

$$\Delta Y_{t+1} = \left(\alpha + \frac{\beta}{\delta} Y_t\right)(\delta - Y_t) \quad (6)$$

Si se opera esta expresión, se obtiene la siguiente formulación:

$$\Delta Y_{t+1} = \alpha\delta - \alpha Y_t + \frac{\beta}{\delta} Y_t \delta - \frac{\beta}{\delta} Y_t^2 \quad (7)$$

Expresión que se puede reescribir como:

$$\Delta Y_{t+1} = \alpha\delta + (\beta - \alpha)Y_t - \frac{\beta}{\delta} Y_t^2 \quad (8)$$

Con el objeto de obtener el punto de saturación, lo que propone el autor es el estimar la ecuación anterior empleando Mínimos Cuadrados Ordinarios; a partir de seguir el siguiente procedimiento:

$$aY_t^2 + bY_t + c = 0 \quad (9)$$

De acuerdo con ello, el valor de saturación δ se estima a partir de las raíces del siguiente polinomio de segundo grado.

$$\delta = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \quad (10)$$

Para poder estimar el proceso se parte de la normalización del proceso de difusión, para ello se calcula el porcentaje respecto del techo estimado.

$$Z_t = \frac{Y_T}{\delta} \quad (11)$$

A partir de este cálculo se obtienen las variaciones de Z_t

$$\Delta Z_t = Z_t - Z_{t-1} \quad (12)$$

La regresión a estimar es la siguiente:

$$\Delta Z_t = \beta Z_t(1 - Z_t) \quad (13)$$

Para completar el proceso de difusión, se toma la expresión anterior adelantada un periodo.

$$\Delta Z_{t+1} = \beta Z_{t+1}(1 - Z_{t+1}) \quad (14)$$

Si reescribimos la expresión anterior podemos obtener la siguiente ecuación de segundo grado en el que la incognita es Z_{t+1}

$$\beta Z_{t+1}^2 + (1 - \beta)Z_{t+1} - Z_t \quad (15)$$

Con lo cual llegamos a la siguiente ecuación que permite predecir los valores que completan el proceso de difusión.

$$Z_{t+1} = \frac{-(1 - \beta) \pm \sqrt{(1 - \beta) + 4\beta Z_t}}{2\beta} \quad (16)$$

2.4. Aproximación a la estimación del Impacto en el Gasto de los Hogares de Lima Metropolitana

Para esta parte se adaptaron los desarrollos de (Maldoom y Horvath, 2002). La metodología planteada por estos autores parte de estimar un diferencial de gasto. Para nuestro caso sería el gasto en GN en el que incurre el hogar i , representado mediante G_{iGN} y el gasto cuando se emplea GLP, G_{iGLP} . El diferencial en el gasto

se especifica a partir de $G_{iGN} - G_{iGLP}$, si la diferencia es negativa implicaría que el hogar gasta menos en GN que en GLP.

Los beneficios que obtiene un hogar producto de emplear GN se pueden agrupar en dos partes: i) un menor gasto de los hogares cuando emplean GN como combustible para cocinar respecto de cuando no emplean este combustible, ii) un conjunto de otros beneficios no observables (como la menor contaminación en el hogar producto de contar con GN). De acuerdo con ello, la ecuación de los beneficios se puede representar de la siguiente manera:

$$\phi_i = \alpha X_i + \epsilon_i \quad (17)$$

En la expresión, ϕ_i representa los beneficios obtenidos por los hogares en el gasto de emplear GN, α es un coeficiente que mide el efecto parcial de algunas características (X_i) asociadas a la elección del GN como combustible para cocinar por parte de los hogares. ϵ_i representa al componente aleatorio que incorpora a las variables no observables que impactarían en el gasto de los hogares.

De este modo, un hogar tomaría la decisión de emplear el GN como combustible para cocinar siempre y cuando el beneficio esperado sea menor que un nivel de reserva, por regularidad se suele asumir que este valor es cero. Para formalizar estas ideas en la siguiente expresión se presenta la diferencia del gasto del hogar cuando emplea el GN, al cual se le agrega un conjunto de beneficios adicionales producto de esta decisión.

$$G_{iGN} - G_{iGLP} + \phi_i < 0 \quad (18)$$

Reemplazando el valor del parámetro ϕ_i se obtiene la siguiente expresión:

$$G_{iGN} - G_{iGLP} + \alpha X_i + \epsilon_i < 0 \quad (19)$$

Entonces, la decisión de emplear GN en el hogar como combustible para cocinar dependerá de que el hogar obtenga beneficios. De este modo, si se define una variable indicadora I_i , que toma el valor de 1 si se emplea GN y 0 si se emplea GLP, se puede representar la decisión de contar con GN mediante la siguiente expresión:

$$I_i = \beta_0 + \beta_1(G_{iGN} - G_{iGLP}) + \beta_2 X_i + \epsilon_i \quad (20)$$

El coeficiente β_0 mide un efecto constante relacionado con la decisión de contar con GN; mientras que β_1 representa el efecto diferencial en el gasto. Finalmente,

en β_2 se incluye el efecto de los controles del modelo, discutidos con anterioridad en la elección del combustible para cocinar.

El problema tras esta formulación es que no se sabe, de manera simultánea, cuál es el gasto del hogar con GN y sin GN; por lo que una manera de aproximar este gasto es mediante la predicción de cocinar con GLP de aquellos hogares que cocinan con GN y cual hubiera sido el gasto en GN de aquellos hogares que emplean GLP.

Para lograr el objetivo, y evitar potenciales problemas de sesgo de selección, vinculados a realizar estimaciones del gasto solamente con aquellos hogares que cocinan con GN, se controla por un factor denominado como la Razón Inversa de Mills. Por lo que la expresión a estimar tendría la siguiente forma:

$$E(G_{iGN}/I_i = 1) = \overline{G_m} + \Gamma \frac{\phi(\Psi)}{1 - \Phi(\Psi)} \quad (21)$$

El factor $\overline{G_m}$ representa el gasto promedio en GN; mientras que ϕ y Φ son la función de densidad y la función acumulada de una normal, respectivamente. De otro lado, Ψ corresponde a la estimación consistente de la probabilidad de que un hogar emplee GN como combustible para cocinar. La estimación del gasto en GLP se presenta la siguiente expresión:

$$E(G_{iGLP}/I_i = 0) = \overline{G_m} + \Gamma \frac{\phi(\Psi)}{\Phi(\Psi)} \quad (22)$$

3. ANÁLISIS Y ESTIMACIONES

3.1. Análisis de la Información

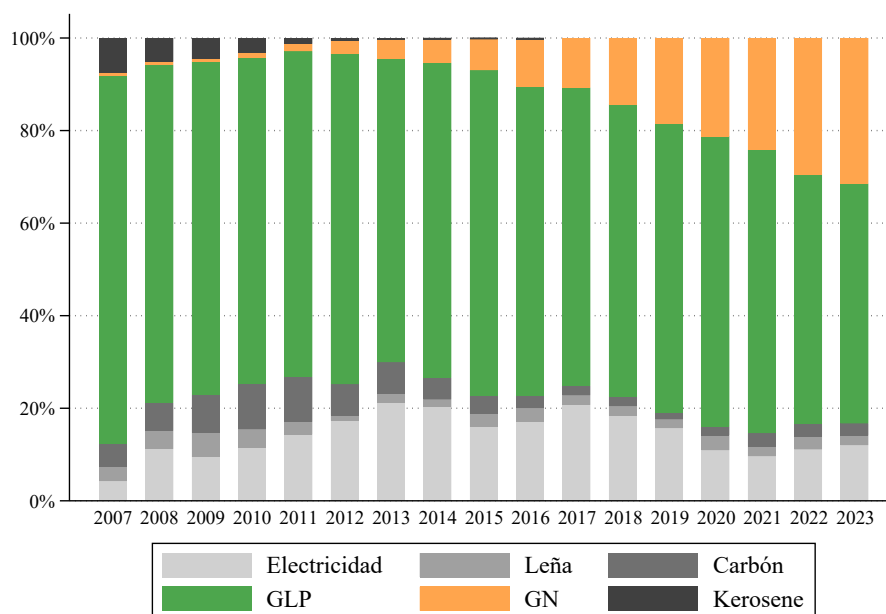
Para la estimación del impacto en el gasto de los hogares, la información con la que se cuenta corresponde a los módulos 15, 34, 100, 200 y 300, de la ENAHO, para los años 2007 al 2023. La información de los módulos fue fusionada para cada uno de los años y se seleccionaron aquellas variables explicativas, que de acuerdo con la literatura revisada, establecen un mecanismo de transmisión con la elección de combustibles para cocinar; además, de aquellas que se relacionan con el gasto de los hogares. De acuerdo con ello, se entiende que existe una decisión de tipo discreta respecto de la elección del combustible para cocinar; y luego de ello, un comportamiento continuo asociado al monto del gasto del hogar en el combustible para cocinar.

De otro lado, para estimar el proceso de difusión del GN en Lima Metropolitana, se cuenta con información de una serie de tiempo de la evolución diaria de la adopción residencial de GN en Lima Metropolitana, desde el 2005 hasta octubre de 2023. Con ello, se pudo identificar que la información obtenida de la adopción residencial del GN en la ENAHO se encuentra subestimada en alrededor de un 60 %; no obstante, la información socioeconómica que se obtiene de los hogares en la ENAHO resulta más que suficiente para poder realizar el ejercicio inferencial acerca del impacto en el gasto de los hogares.

Respecto de la elección del combustible para cocinar en Lima Metropolitana, en el siguiente gráfico se presenta la evolución del portafolio de combustibles usados para cocinar en Lima Metropolitana para el periodo 2007 al 2023. De esta evolución podemos concluir los siguientes hechos estilizados: i) el GLP es el combustible empleado con mayor frecuencia en las actividades de cocción de alimentos por parte de los hogares de Lima Metropolitana, ii) la disminución en el uso del kerosene ⁹, iii) el incremento sostenido del GN, a partir del 2010, como combustible para cocinar por parte de los hogares de Lima Metropolitana.

⁹A partir del 2009 se prohibió la venta de Kerosene y se desarrolló un programa de sustitución del kerosene por el GLP

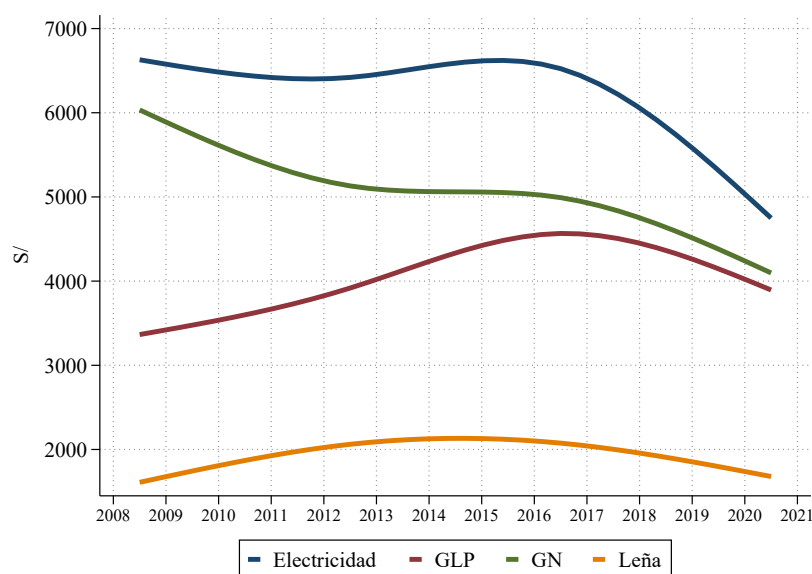
Figura 1: Evolución de los Combustibles para Cocinar en Lima Metropolitana (2007-2023)



Fuente: ENAHO 2004-2022, Elaboración: Propia

Una de las variables claves discutida, a lo largo de esta investigación, es el ingreso, según los datos de la ENAHO, en Lima Metropolitana los hogares con el ingreso promedio mensual más alto prefieren usar electricidad en sus actividades de cocción de alimentos. En segundo lugar, el promedio más alto corresponde a los hogares que emplean GLP; un detalle a notar es que, en los últimos años, el promedio de los ingresos de los hogares que usan el GN se encuentra muy cercano al del GLP. Finalmente, los hogares con el ingreso promedio más bajo emplean leña como combustible para cocinar.

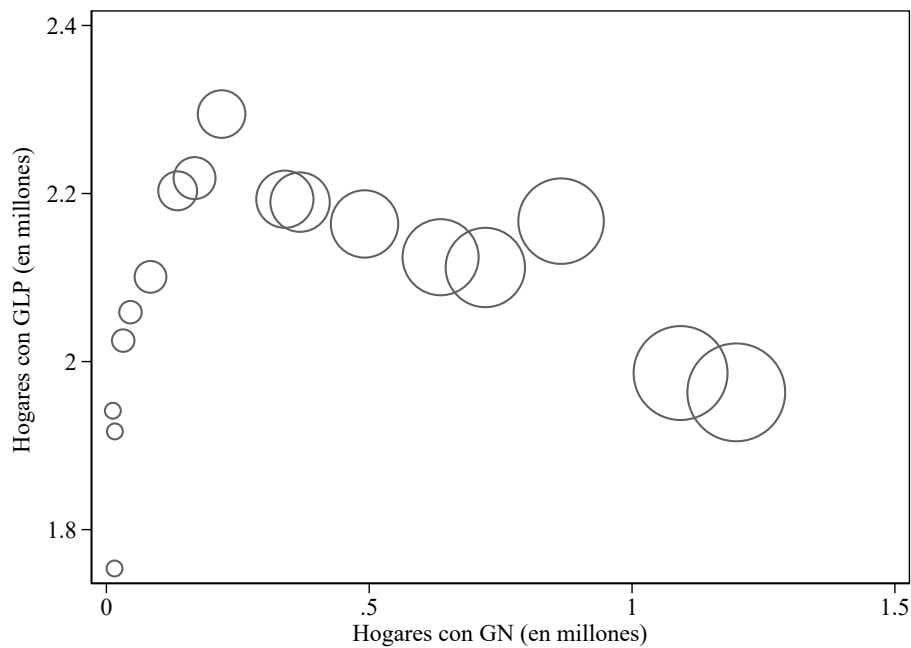
Figura 2: Ingresos de los Hogares y Combustibles que Emplean para Cocinar



El modulo 100 de la ENAHO permite identificar al total de hogares que emplean GLP y GN como combustible principal para la cocción de alimentos. En el siguiente gráfico podemos observar que inicialmente existía una relación directa entre el total de hogares que emplean GLP para cocinar y los que usan GN, relación que se volvió inversa. Complementando ello, los hogares que emplean GLP para cocinar presentan una tendencia decreciente, pasando de 99 % en el 2007 a 62 % en el 2023; mientras que con el GN ocurre lo contrario, el porcentaje de hogares paso de 1 % a 38 % en los años señalados.¹⁰

¹⁰Cabe señalar que los porcentajes obtenidos no toman en cuenta aquellos hogares que emplean combustibles para cocinar diferentes al GLP y al GN.

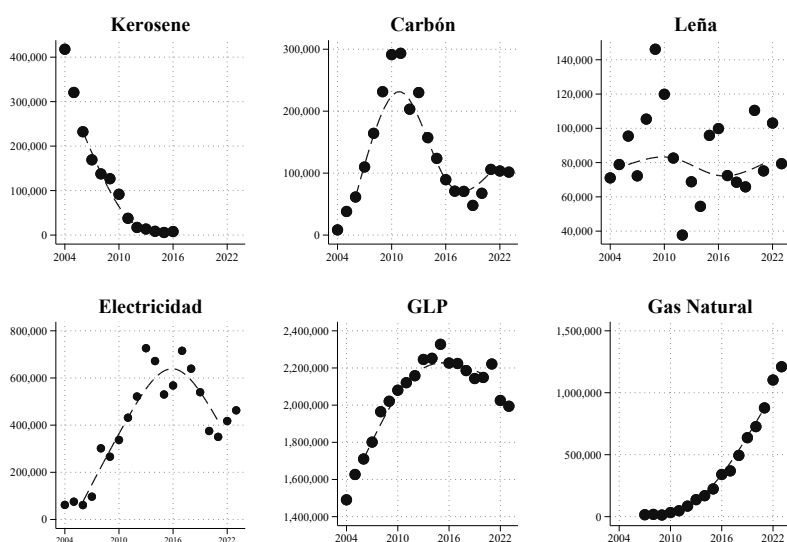
Figura 3: **Correlación entre los Hogares que Emplean GN y GLP**



3.2. Estimación del Proceso de Difusión del GN en Lima Metropolitana

De acuerdo con la especificación del modelo de difusión presentado previamente, en el siguiente gráfico se presenta el proceso de adopción de los seis combustibles para cocinar empleados en Lima Metropolitana, en el periodo 2004 a 2023. Respecto de estos, podemos identificar que un conjunto de combustibles parecen haber completado su proceso de difusión, como el carbón el cual se encontraría en una fase de desadopción del servicio, en esa misma línea se encontraría la leña, aunque con un cierto incremento en la adopción en los últimos años. En el caso de la electricidad y el GLP, estas ya habrían ingresado en la fase de desadopción, concerniente al kerosene, se prohibió su uso a partir del 2010; por lo que su tasa de adopción resulta de cero a partir de años posteriores. El combustible que viene mostrando una tendencia creciente es el GN.

Figura 4: Procesos de Difusión de Combustibles para Cocinar en Lima Metropolitana



Fuente: ENAHO 2004-2022, Elaboración: Propia

Como punto de partida de la estimación de un proceso de difusión se requiere el contar con el techo del mismo. Para esta estimación se empleó información de una serie de tiempo de 6282 datos acumulados de hogares que cuentan con GN como combustible para cocinar, correspondiente al 14 de marzo de 2005 al 1 de octubre de 2023. Con estos datos se obtuvo un techo de 3000579 hogares; es decir, se espera que el proceso de adopción del GN en Lima Metropolitana alcance este valor máximo en número de hogares que emplean GN. Podemos ver a continuación los resultados de la estimación de la ecuación que permitió estimar el techo¹¹.

Cuadro 1: Regresión para Obtener el Techo

	Estimación del Techo		
	b	t	p
Hogares con GN al Cuadrado	-0.00	-24.10	0.00
Hogares con GN	0.00	50.75	0.00
Constant	26.21	7.18	0.00
Observaciones	6281.00		
R2 Ajustado	0.59		
Prueba F	4602.41***		

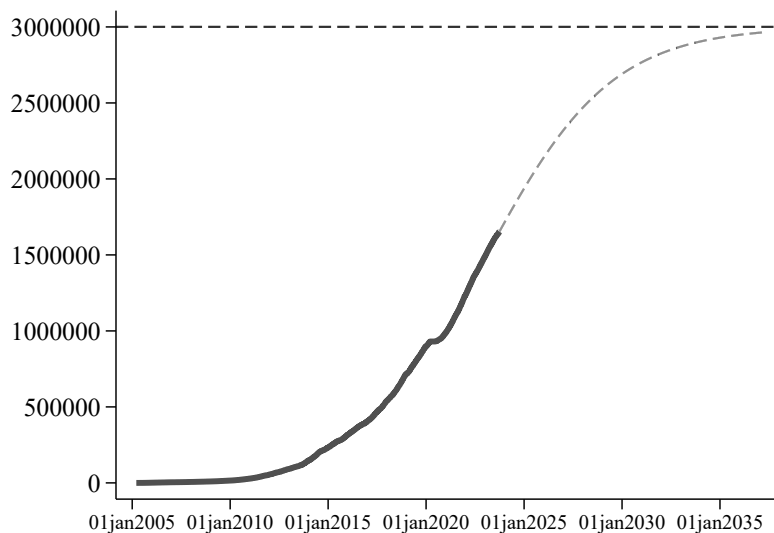
Luego de la estimación del techo del proceso de difusión, se determinó la fase en la cual se encuentra el proceso de adopción y se proyectaron las siguientes fases.

¹¹Para estimar el techo se elaboró el programa `regR`

De acuerdo con (Rivas, 2009) la fase de los **innovadores** representan el 2.5 % de los adoptantes de un nuevo producto; luego de ello, los **adoptantes tempranos** se encuentra representada por el 13.5 % de los adoptantes. La siguiente fase esta representada por la **mayoría temprana** la cuál se alcanza en el punto de máximo crecimiento de la adopción, **punto de inflexión**, la cual ocurre con el 50 % de los adoptantes; porsterior a ello vienen los adoptantes de la **mayoría tardia**, los cuales acumulan al 84 % de los adoptantes; finalmente, la última fase estaría compuesta por los **adoptantes rezagados**.

Respecto a la estimación realizada, se obtuvo que la fase de **innovadores** duró desde el lanzamiento del GN hasta mayo del 2008 con un estimado de 75014 adoptantes. La fase de **adoptantes tempranos** duró hasta el 23 de junio de 2017 con un estimado de 465089 adoptantes. Fase que duró hasta el **punto de inflexión**, el cuál ocurrió en enero de 2023. Luego de ello, la fase de la **mayoría tardia** se espera culmine en mayo del 2028, con un estimado de 2520486 hogares que cocinan con GN; finalmente, los **adoptantes rezagados** estarían conformados por aquellos hogares que adoptan el GN hasta que se alcance el punto de saturación de 3000579 de hogares, el cuál se estima que ocurra a fines del 2034.

Figura 5: Procesos de Difusión de Combustibles para Cocinar en Lima Metropolitana



Fuente: OSINERGMIN, Elaboración: Propia

De esta parte se concluye que la adopción del GN como combustible para cocinar en Lima Metropolitana se encontraría en la etapa de la mayoría tardía; de acuerdo con ello, el impacto a estimar se asocia con las fases descritas.

3.3. Estimación del Impacto

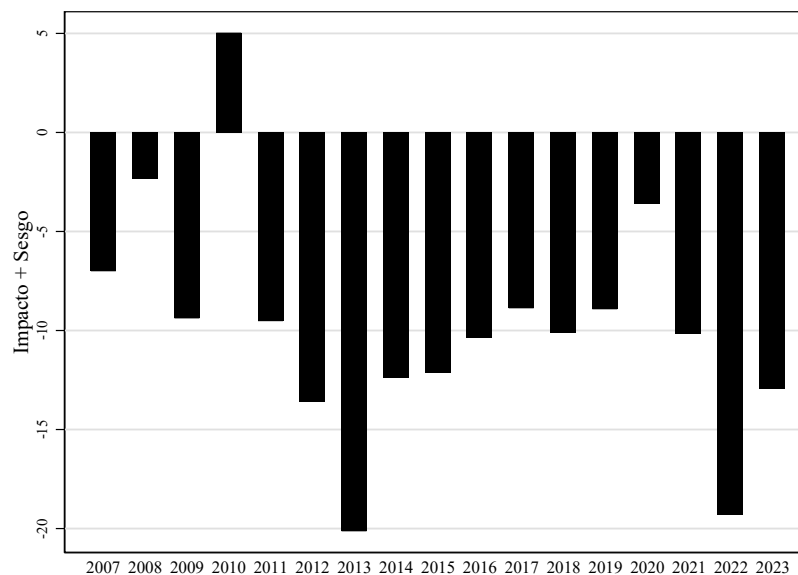
En esta sección se estimará el impacto en el gasto del hogar del empleo del GN, comparada con el empleo de GLP.

3.3.1. Diferencia de Medias

El efecto a estimar tiene que ver con la diferencia promedio en el gasto de aquellos hogares que emplean GN como combustible para cocinar y los que emplean GLP. Un resultado negativo, y estadísticamente significativo, significa que emplear GN trae como consecuencia una reducción en el gasto del hogar.

Para evaluar este efecto, se plantea una estimación inicial que será comparada con los resultados de las diferentes especificaciones que se ensayan en este documento. Para ello, se realiza una regresión simple, mediante Mínimos Cuadrados Ordinarios, entre la variable gasto y una variable dummy que toma el valor de 1 si el hogar emplea GN y 0 si emplea GLP. Los resultados se presentan a continuación para cada uno de los 17 años de estudio. Tal como se observa, existe un impacto, la mayoría de las veces, a favor del ahorro de los hogares que emplean GN. Con la única excepción el año 2010 donde el impacto estimado indica que los hogares que emplearon GLP ahorraron más que aquellos que emplearon GN. El valor promedio del impacto en el ahorro de los hogares fue de S/ 9.8; mientras que los mayores impactos se obtuvieron en el año 2013 y en el 2022 (arlededor de S/20).

Figura 6: Test de Medias



Tal como se puede apreciar del 3 se realizaron 17 regresiones, una por cada año, para más de 1 700 000 hogares de Lima Metropolitana. Para cada uno de los años el coeficiente de la variable `dummy`, que refiere a sí el hogar emplea GN o GLP como combustible para cocinar ($\bar{y}_1 - \bar{y}_0$), resultó estadísticamente significativo. No obstante, dado que la asignación de los hogares en la decisión del empleo del combustible para cocinar no resulta aleatoria, sino a partir de que se cumplan ciertas características (principalmente que la red de GN se encuentre cercana a la vivienda), se debe notar que el resultado estimado contiene al **Efecto Promedio en el Tratamiento (ATE)** y el **sesgo de selección producto de esta autoselección**.

$$\bar{y}_1 - \bar{y}_0 = ATE + \text{Sesgo de Selección}$$

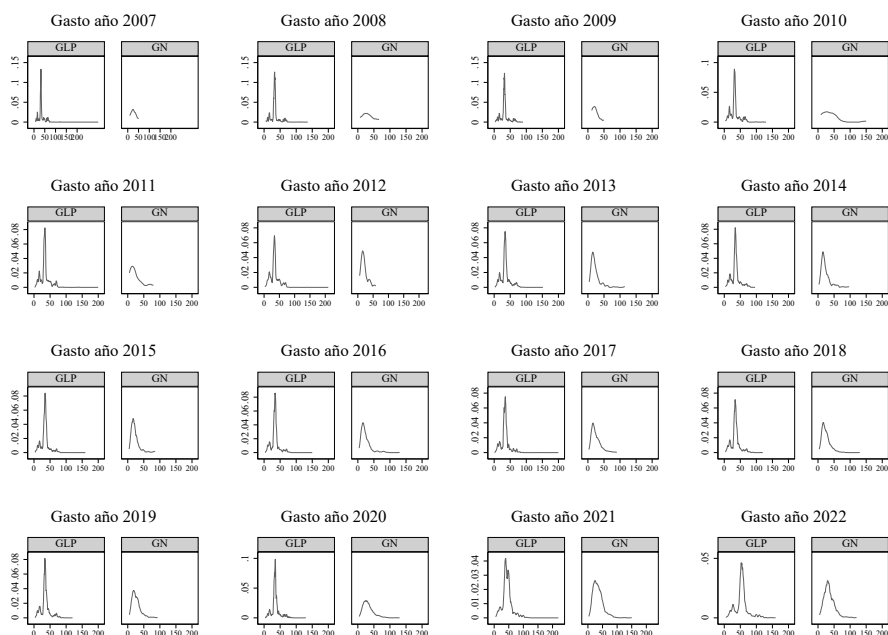
De acuerdo con los señalado, volvemos a realizar la regresión pero ahora incluyendo algunas variables de control disponibles en la base de datos y referidas al pretratamiento. Para la elección de las variables se toma en cuenta la revisión de literatura realizada con anterioridad; además de los siguientes 3 factores relevantes recomendados por (Rivas, 2020) vinculados con la estimación de la probabilidad de la elección del combustible para cocinar: i) mayor cantidad de combustibles para cocinar involucran un mayor número de parámetros que se deben estimar, ii) existe heterogeneidad en los gustos de los consumidores, iii) problemas de endogeneidad típicos de no contar con la variable precio.

Para nuestro caso la variable dependiente es el gasto en combustibles para cocinar tanto para los hogares que emplean GN como aquellos que usan GLP; por lo que no se tiene el mismo problema de contar con múltiples parámetros que estimar como en el caso del modelo Logit Multinomial estimado por (Rivas, 2020). Un problema asociado sería el de la distribución del gasto entre hogares que emplean GLP y GN; al respecto, (Osinermin, 2021) empleando la Encuesta Residencial de Consumo y Usos de Energía para los años 2019-2020 estimó que en Lima Metropolitana, en promedio, se consumen 1.01 balones de GLP por mes; por lo que el valor esperado por el gasto en este concepto para un hogar no podría superar los S/ 150 soles (algó más de 2 balones de GLP). Sin embargo, en la distribución del gasto se observan comportamientos atípicos por encima de este valor, que podrían reflejar que existen hogares que emplean el balón de GLP no solamente para cocinar sino también para otras actividades, como las comerciales. En el caso del GN, para ciertos años también se evidencia este comportamiento, lo que indicaría es que podrían estar empleando este combustible para más de un punto de conexión; por ejemplo para conectarlo a la terma o la lavadora. Esto nos lleva a pensar que para realizar la estimación del impacto requerimos el determinar un **soporte común**.

En el siguiente gráfico 7 podemos observar la evolución de la distribución del gasto tanto en GLO como en GN para los 17 años de estudio. Un detalle a notar es que el gasto en GLP muestra una distribución con una cola hacia la derecha y lo mismo ocurre con el GN para ciertos años. Este comportamiento muestra que existen **valores atípicos** que estarían influyendo en los resultados de la estimación realizada; además, que podrían estar referidas a actividades distintas

del hogar de las de cocinar.

Figura 7: Distribución del Gasto por Combustibles para Cocinar en Lima Metropolitana



Resulta cierto que para realizar la estimación se está trabajando con datos observacionales. En ese sentido (Coma, 2012) señala que la asignación aleatoria al tratamiento se considera el estándar de oro; pues las asignaciones a los grupos de tratamiento y control deberían ser estadísticamente equivalentes en características observables y no observables. Por lo que al no estar en dicho escenario se requerirá, en primer lugar, de la inclusión de variables de control de pre-tratamiento que nos ayuden a aislar el efecto del programa en las variables de interés.

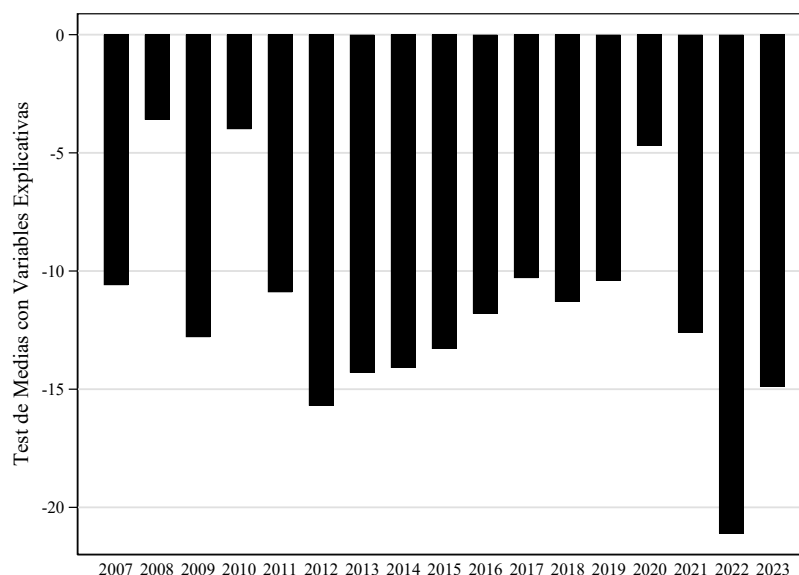
Respecto de que variables de pre-tratamiento a incluir, una primera discusión tiene que ver con la heterogeneidad en los gustos de los consumidores de Lima Metropolitana; al respecto, se debe señalar que ha existido una migración importante de personas desde el campo a la ciudad, los cuales trajeron sus costumbres en comidas y preparación; sin embargo, a diferencia de algunos lugares del interior donde se emplea principalmente la leña para cocinar, en Lima Metropolitana ha existido una convergencia en la elección de combustibles para cocinar, y tal como se mostró con anterioridad los combustibles que se emplean con mayor frecuencia son el GLP y el GN. Entonces, para Lima Metropolitana la heterogeneidad no se encontraría en función de las costumbres asociadas a la procedencia de las personas.

Concerniente a la selección de las variables de pre-tratamiento, se tomará como

referencia los desarrollos de (Soberon-Ferrer y Dardis, 1991) quienes identifican variables vinculadas con la **Producción de los Hogares**, las cuales tienen que ver con los ingresos que genera el hogar y que le permiten cubrir los gastos asociados. Luego de ello, se tienen variables del *Ciclo de Vida del Hogar*, estas incluyen a la edad del jefe del hogar, el número de niños menores de 6 años, el número de personas que perciben ingresos en el hogar, variables que son importantes para explicar la diferencia en gastos entre hogares (Bellante y Foster, 1984). **Educación del jefe del hogar** variable que se relaciona con la eficiencia en el gasto del hogar. Otra variable asociada es **Ubicación**, pues es de esperar que familias que viven en áreas urbanas gasten más que aquellas que viven en áreas rurales; al existir una mayor probabilidad, por ejemplo, que en el área urbana paguen por ciertos servicios domésticos comparado con el área rural en el que los propios hogares desarrollan estos servicios. *Propiedad de la casa* pues el gasto de los que alquilan suele ser menor que el de los propietarios de un inmueble.

A partir de este conjunto de variables identificadas, además de las modificaciones señaladas a la base de datos, se procede a estimar nuevamente el modelo de la diferencia de gastos incluyendo variables de pre-tratamiento. Los resultados se presentan en el siguiente gráfico evolutivo 8 y muestran que el ahorro promedio de los hogares que cuentan con GN sería de S/ 11.55; mientras que el mayor impacto, el cual supera los S/ 20, se habría obtenido el 2022.

Figura 8: Test de Medias Empleando Variables Explicativas



Los resultados de las estimaciones se presentan en el anexo 4 y se identifica que todos los coeficientes resultaron estadísticamente significativos y obtuvieron el signo esperado. Además, la variable de gasto en GN tuvo un signo negativo para todos los años considerados.

No obstante, las estimaciones obtenidas hasta el momento resultarían consistentes si es que el gasto de los hogares con GN es independientes de los resultados potenciales; es decir: i) el valor promedio de las variables con las que se esta midiendo el impacto, para los hogares que cuentan con GN y con GLP, debería ser el mismo, ii) del mismo modo, el gasto promedio de los hogares con GN y con GLP debería ser el mismo si ambos grupos consumieran GN. En ausencia de cumplimiento de estas condiciones de independencia de los resultados potenciales se podría tener un problema de sesgo de selección (Cameron y Trivedi, 2005)

Lo cierto es que la forma más efectiva de que se cumpla la condición de independencia de resultados potenciales es mediante asignaciones aleatorias de las empresas a cada uno de los grupos (Cameron y Trivedi, 2005)

Toda esta discusión nos conduce a lo señalado por (Holland, 1986) quien afirma que el problema fundamental de la inferencia causal es que es imposible que podamos observar los resultados obtenidos por un individuo que fue beneficiado y a la vez cuales serían los beneficios que habría obtenido si este no hubiera formado parte del programa. Problema que es persistente debido a que los datos que enfrentamos resultan observacionales.

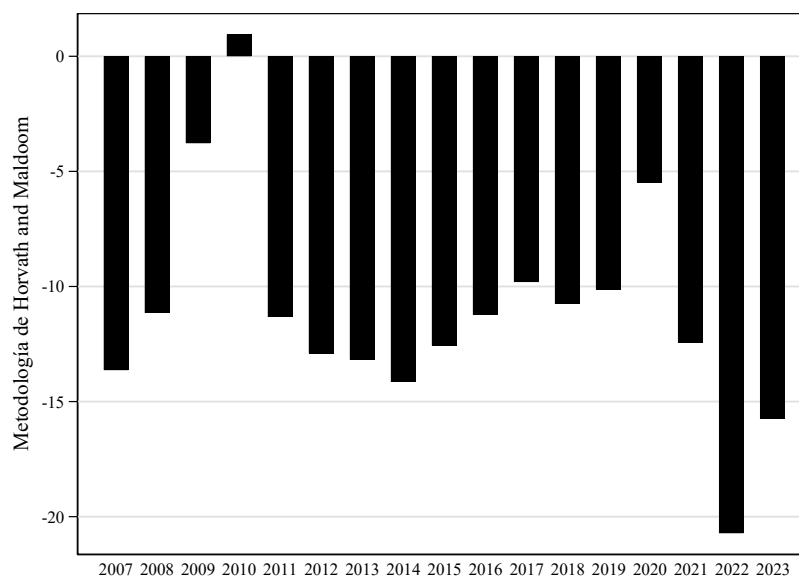
Dado que estamos claro que la asignación en el tratamiento no resulta aleatoria; por lo que depende de un conjunto de variables observables. En ese contexto el concepto de **Propensity Score Matching** resulta útil. No obstante ello, antes exploraremos los resultados que se obtendrían con la metodología descrita con anterioridad de (Maldoom y Horvath, 2002).

3.3.2. Estimación Mediante el Método de Horváth y Maldoom

Como se señaló, esta metodología estima el proceso en dos etapas de elección del combustible para cocinar y; condicionado a ello, los determinantes del gasto para cocinar. En la versión desarrollada por los autores (Maldoom y Horvath, 2002) se emplea para determinar efectos de sustitución por el lado del acceso y uso entre la telefonía móvil y el teléfono fijo. En la adaptación de esta metodología se emplea para pronosticar cual habría sido el gasto en GN de aquellos hogares que emplean GLP y a partir de dichos resultados estimar un test de medias para determinar el impacto en el gasto.

Los resultados de la aplicación de esta metodología se presentan en el siguiente gráfico 9 evolutivo y el impacto promedio obtenido fue de S/ 11.06 de ahorro para los hogares que emplean GN como combustible principal para cocinar sus alimentos; además se volvió a obtener que el 2022 fue el año en el que se obtuvo un mayor impacto y este resultó mayor a S/20. Se debe notar que el 2010 el resultado que se obtuvo fue positivo similar al test de medias inicial que se realizó.

Figura 9: Estimación a partir del Procedimiento de Horváth y Maldoom



Los resultados de las estimaciones se presentan en los cuadros 4 5 y 5 7 del Anexo. Podemos observar de los resultados que la Razón Inversa de Mills fueron estadísticamente significativos en el 2012 y a partir del 2015 en adelante.

No obstante, con esta técnica aún no se resuelve el problema de emplear datos observacionales. Por ello, la idea del Propensity Score Matching es hacerle frente a los problemas de sesgo de selección, propios de selecciones no alaeatorias, a partir del desarrollo de métodos de emparejamiento (matching) que tratan de identificar unidades no tratadas similares en características observadas a una unidad participante, con ello podemos tener una diferencia de medias del impacto del gasto promedio de los hogares minimizando los problemas de sesgo.

3.3.3. Estimación Mediante Propensity Score Matching

En esta sección, se hará frente a los problemas de aleatoriedad propio de trabajar con datos observacionales, para resolver ello se empleará Propensity Score Matching. No obstante, uno de los problemas que se enfrenta con esta técnica es el de dimensionalidad, al existir múltiples variables que se podrían usar; al respecto, (Rosenbaum y Rubin, 1983) resolvieron este problema a partir de una métrica única denominada Propensity Score a partir de la estimación de un modelo logit o probit.

En el caso de la aplicación del Propensity Score Matching se requiere que se cumpla el supuesto de independencia condicional y de soporte común, pues a partir de

ello se puede reducir el sesgo de selección vinculado con características observables (Coma, 2012). El supuesto de independencia condicional señala que los resultados potenciales no dependen del tratamiento, también se le denomina selección en observables, lo que implica que las características no observables no influyen en la determinación de la participación (Dehejia y Wahba, 2002). Para que se cumpla dicho supuesto, se requiere el controlar por todas las características observadas de tal modo que las unidades de comparación deberían tener en promedio el mismo resultado que las de tratamiento en ausencia de intervención.

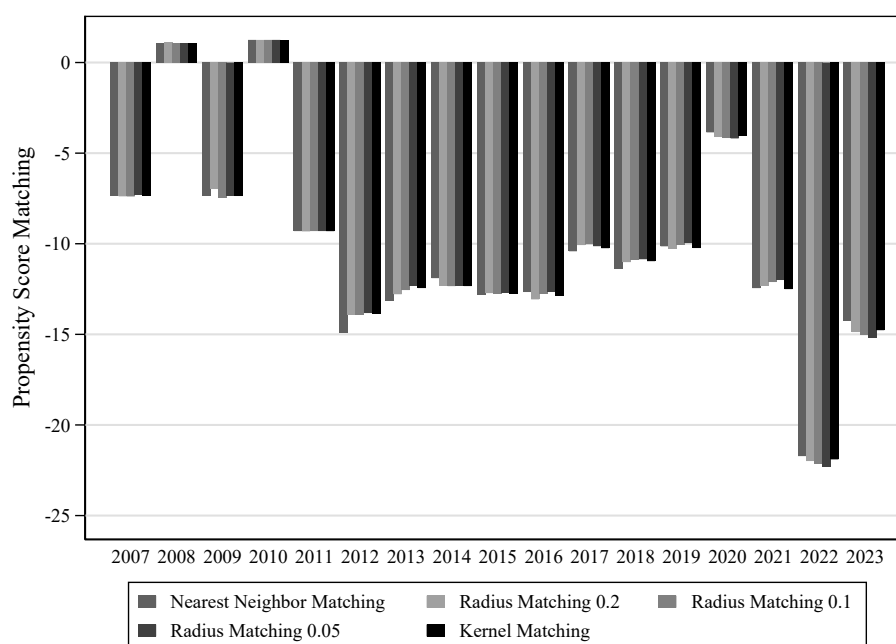
La hipótesis de soporte común indica que existe un solapamiento en la distribución de observables entre ambos grupos (Coma, 2012). Para la estimación de un PSM sin ponderar se emplea el comando `pselect`, ello en función del trabajo de (Imbens y Rubin, 2015). Además, se incluyeron como variables observables aquellas que permiten explicar el gasto de los hogares. El comando `pselect` seleccionó a todas estas variables para ser incluidas en la especificación.

Respecto a la elección del algoritmo de matching, el más usado es el del vecino más cercano, el cual selecciona a al grupo de control en función de la distancia (comandos de STATA: `attnd`, `attnw`). También se cuenta con los algoritmos de radio que selecciona a los controles a partir de cierta distancia predefinida (comandos de STATA: `attr`, `atts`). Además, se puede dividir el soporte común en estratos o el emplear kernels (comando en STATA: `atrk`), el cual toma el promedio ponderado de todas las observaciones en el soporte común. (Maffioli, 2010)

Los algoritmos seleccionados para estimar el impacto fueron aquellos que permitieron la inclusión del **factor de expansión**, que es el ponderador de cada uno de los hogares considerados en la muestra. De acuerdo con ello se empleó **Nearest Neighbor Matching**, **Radius Matching** y **Kernel Matching**. Estos algoritmos emplean diferentes criterios para seleccionar a los grupos de control. Por ejemplo, el algoritmo del vecino más cercano "Nearest Neighbor Matching" lo que hace es ordenar la base de datos, según el score generado, y con ello busca los datos más próximos comparando hacia arriba y hacia abajo. En el caso del matching empleando un radio, se considera una distancia dentro del cual se definen los emparejamientos. Finalmente, Kernel Matching asigna un peso a cada unidad de control en función de su distancia del puntaje de la unidad tratada. Las unidades más cercanas reciben mayor peso y las más lejanas menor peso.

Los algoritmos empleados nos brindan resultados similares del impacto, con un ahorro promedio de $S/10$. El 2022, sigue siendo el año en el que se observó un mayor impacto; mientras que los años 2008 y 2010 muestran que el ahorro en GLP fue mayor al del GN; no obstante, para ninguno de los algoritmos empleados el resultado fue estadísticamente significativo.

Figura 10: Estimación Empleando Propensity Score Matching



Los resultados de las estimaciones realizadas se pueden observar en el cuadro 6 del anexo. Respecto de la validez de los mismos, (Maffioli, 2010) propone los siguientes criterios para evaluar la calidad del matching: i) chequear la reducción del sesgo estandarizado antes y después del matching, ii) testear la igualdad de medias en los grupos de tratados y de control, antes y después del matching (uso del comando `pstest`), iii) chequear la significancia conjunta del modelo de participación, después del matching se debe tener un pseudo-R2 bajo (uso del comando `pstest`). Para el caso del soporte común, el autor propone el emplear el test "two-sample Kolmogorov-Smirnov". Respecto al primer punto, tomando en cuenta una base de datos comparable, el test de medias que incluye variables de pre-tratamiento nos dió un impacto promedio de $S/ 11.55$; mientras que el procedimiento de Horváth y Maldoom, de $S/ 11.06$, valores que se encuentran por encima del promedio obtenido mediante Propensity Score Matching.

Un tema de discusión es que si bien Propensity Score Matching elimina el sesgo asociado a las características que se observan vinculadas con el empleo del GN como combustible para cocinar, pero pueden existir características temporales que son constantes y que afecten la decisión del empleo del GN (como el crecimiento de la red de ductos en Lima Metropolitana). Este último factor solo se podría recoger mediante el método de Diferencias en Diferencias. Al respecto, (Maffioli, 2010) recomienda combinar ambos métodos en busca de mejorar las estimaciones. Para este caso, esto se convierte en una limitación pues se cuenta con 17 bases de datos de corte transversal; pero estas no constiuyen un panel de datos.

4. Resultados Finales

Para obtener el impacto final, en términos de ahorros agregados y según la fase de adopción del GN, se toma un promedio de los impactos obtenidos con cada una de los algoritmos de Propensity Score Matching, exceptuando los años 2008 y 2010 en el que los resultados fueron positivos pero no estadísticamente significativos. De acuerdo con ello, en el siguiente cuadro podemos notar que el mayor ahorro que obtuvieron los hogares de Lima Metropolitana por usar GN fue en la mayoría temprana con cerca de 14 millones de soles. En el cuadro también se presenta los ahorros ajustados, los cuales se obtuvieron a partir de asumir que los hogares ahorran ese dinero en un banco ¹². Con ello se pudo determinar que el ahorro total obtenido por los hogares de Lima Metropolitana fue del orden de los S/ 24 254 562.

Cuadro 2: Impacto Según Fase

Fase	S/	S/ Ajustados
Innovadores	14 629.8	23 682
Adoptantes Tempranos	5 313 790	6 908 687
Mayoría Temprana	13 929 967	15 149 945
Mayoría Tardia	2 141 293	2 172 248
Total		24 254 562

¹²La información de la tasa de interés pasiva se obtuvo del siguiente link:
<https://www.sbs.gob.pe/app/pp/EstadisticasSAEEPportal/Paginas/TIPasivaMercado.aspx?tip=B>

5. Conclusiones

Al 2023, el total de hogares que empleaban Gas Natural como insumo para cocción de alimentos en Lima Metropolitana, se multiplicó por 78, si se le compara con el 2007; mientras que para el mismo periodo de tiempo, el total de hogares que emplean el balón de Gas Licuado de Petróleo se multiplicó solamente por 1.1. Resultado que implica la existencia de un proceso de sustitución entre ambos combustibles.

Para cumplir el objetivo, se empleó información socioeconómica de 17 cortes transversales provenientes de la Encuesta Nacional de Hogares correspondiente al periodo 2007 al 2023. Con esta información se estimaron 7 especificaciones del impacto monetario para cada uno de los años, obteniéndose resultados corregidos con *Propensity Score Matching*. Además, mediante información real de la evolución de la adopción del Gas Natural en Lima Metropolitana, concluyéndose que nos encontraríamos en el periodo de **mayoría tardía**

De las estimaciones realizadas se determinaron impactos diferenciados según las fases del proceso de difusión, determinándose que el mayor impacto se obtuvo en la fase de **Mayoría Temprana**. Respecto al impacto agregado este se estimó en S/ 24 254 562 de menor gasto para aquellos hogares que emplean GN en lugar de GLP.

Referencias

- Akpalu, W., Dasmani, I., y Aglobitse, P. B. (2011). Demand for cooking fuels in a developing country: To what extent do taste and preferences matter? *Energy Policy*, 39(10), 6525–6531.
- Alvis-Guzmán, N., Alvis-Estrada, L., y Orozco-Africano, J. (2008). Costo efectividad del gas natural domiciliario como tecnología sanitaria en localidades rurales del caribe colombiano. *Revista de Salud Pública*, 10, 537–549.
- Barnes, D. F., Krutilla, K., y Hyde, W. F. (2010). *The urban household energy transition: social and environmental impacts in the developing world*. Routledge.
- Bass, F. M. (1969). A new product growth for model consumer durables. *Management science*, 15(5), 215–227.
- Becker, G. S. (1965). A theory of the allocation of time. *The economic journal*, 75(299), 493–517.
- Bellante, D., y Foster, A. C. (1984). Working wives and expenditure on services. *Journal of Consumer Research*, 11(2), 700–707.
- Berry, S. T., Levinsohn, J. A., y Pakes, A. (1993). *Automobile prices in market equilibrium: Part i and ii*. National Bureau of Economic Research Cambridge, Mass., USA.
- Cameron, A. C., y Trivedi, P. K. (2005). *Microeconometrics: methods and applications*. Cambridge university press.
- Coma, M. R. (2012). Técnicas de evaluación de impacto: propensity score matching y aplicaciones prácticas con stata. *Documentos-Instituto de Estudios Fiscales*(2), 1–58.
- Dehejia, R. H., y Wahba, S. (2002). Propensity score-matching methods for nonexperimental causal studies. *Review of Economics and statistics*, 84(1), 151–161.
- Duflo, E., Greenstone, M., y Hanna, R. (2008). Indoor air pollution, health and economic well-being. *SAPI EN. S. Surveys and Perspectives Integrating Environment and Society*(1.1).
- Geist, H. J., y Lambin, E. F. (2002). Proximate causes and underlying driving forces of tropical deforestation: Tropical forests are disappearing as the result of many pressures, both local and regional, acting in various combinations in different geographical locations. *BioScience*, 52(2), 143–150.
- Giovanis, A. N., y Skiadas, C. H. (2007). A new modeling approach investigating the diffusion speed of mobile telecommunication

- services in eu-15. *Computational economics*, 29, 97–106.
- Gupta, G., y Köhlin, G. (2006). Preferences for domestic fuel: analysis with socio-economic factors and rankings in kolkata, india. *Ecological Economics*, 57(1), 107–121.
- Heltberg, R. (2004). Fuel switching: evidence from eight developing countries. *Energy economics*, 26(5), 869–887.
- Hiemstra-Van der Horst, G., y Hovorka, A. J. (2008). Reassessing the “energy ladder”: Household energy use in maun, botswana. *Energy Policy*, 36(9), 3333–3344.
- Holland, P. W. (1986). Statistics and causal inference. *Journal of the American statistical Association*, 81(396), 945–960.
- Hosier, R. H., y Dowd, J. (1987). Household fuel choice in zimbabwe: an empirical test of the energy ladder hypothesis. *Resources and energy*, 9(4), 347–361.
- Imbens, G. W., y Rubin, D. B. (2015). *Causal inference in statistics, social, and biomedical sciences*. Cambridge University Press.
- Kroon, B. v. d., Brouwer, R., y Beukering, P. (2013). The energy ladder: Theoretical myth or empirical truth? results from a meta-analysis. *Renewable and Sustainable Energy Reviews*, 20(504-513), 2019.
- Maffioli, A. (2010). Métodos no experimentales: Propensity score matching. *Curso de Evaluación de Impactos, Banco Interamericano de Desarrollo (bid), Buenos Aires y La Plata, Argentina: Ministerio de Economía de la Provincia de Buenos Aires*.
- Maldoom, D., y Horvath, R. (2002). Fixed mobile substitution: a simultaneous equation model with qualitative and limited dependent variables. En *regional its conference in madrid*.
- Masera, O. R., Saatkamp, B. D., y Kammen, D. M. (2000). From linear fuel switching to multiple cooking strategies: a critique and alternative to the energy ladder model. *World development*, 28(12), 2083–2103.
- Mekonnen, A., y Köhlin, G. (2009). Determinants of household fuel choice in major cities in ethiopia.
- Ngui, D., Mutua, J., Osiolo, H., y Aligula, E. (2011). Household energy demand in kenya: An application of the linear approximate almost ideal demand system (la-aids). *Energy policy*, 39(11), 7084–7094.
- Osinerghmin. (2018). REPORTE DE ANÁLISIS ECONÓMICO SECTORIAL SECTOR GAS NATURAL.

- Osinermin. (2021). Informe de Resultados Consumo y Usos de los Hidrocarburos Líquidos y GLP.
- Ouedraogo, B. (2006). Household energy preferences for cooking in urban ouagadougou, burkina faso. *Energy policy*, 34(18), 3787–3795.
- Rao, M. N., y Reddy, B. S. (2007). Variations in energy use by indian households: An analysis of micro level data. *Energy*, 32(2), 143–153.
- Rivas, J. M. (2009). Modelo de Difusión Tecnológica: Un Análisis de la Industria Peruana de Servicios Móviles.
- Rivas, J. M. (2020). Elección del gas natural como combustible para cocinar en lima metropolitana. *Revista de Análisis Económico y Financiero*, 2(2).
- Rosenbaum, P. R., y Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41–55.
- Soberon-Ferrer, H., y Dardis, R. (1991). Determinants of household expenditures for services. *Journal of consumer research*, 17(4), 385–397.
- Torres, H., Agreda, J., y Polo, C. (2010). Impacto ambiental producido por el uso de lena en el área de conservación regional vilacota-maure de la región tacna. *Universidad Nacional Jorge Basadre Grohmann-Peru*.
- Van der Kroon, B., Brouwer, R., y Van Beukering, P. J. (2013). The energy ladder: Theoretical myth or empirical truth? results from a meta-analysis. *Renewable and sustainable energy reviews*, 20, 504–513.

6. ANEXOS

Cuadro 3: Regresión Efecto Agregado

	G2007	G2008	G2009	G2010	G2011	G2012	G2013	G2014	G2015	G2016	G2017	G2018	G2019	G2020	G2021	G2022	G2023
Prob(GN=1)	-7.01***	-2.35***	-9.39***	5.04***	-9.53***	-13.61***	-20.14***	-12.39***	-12.14***	-10.36***	-8.88***	-10.10***	-8.93***	-3.61***	-10.16***	-19.30***	-12.94***
	(0.12)	(0.11)	(0.11)	(0.08)	(0.07)	(0.05)	(0.45)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Constant	34.40***	33.85***	32.48***	32.98***	34.36***	34.43***	42.63***	35.20***	34.43***	34.11***	33.86***	34.82***	35.24***	35.86***	44.78***	54.30***	47.14***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.11)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Observations	1754176	1898582	1922472	2012000	2062375	2146647	2302293	2347182	2481166	2485031	2518751	2592436	2708828	2787868	2970014	2988619	2999173
R^2	0.002	0.000	0.004	0.002	0.009	0.030	0.001	0.052	0.067	0.069	0.053	0.078	0.071	0.013	0.068	0.193	0.111

Standard errors in parentheses

* $p < 0,05$, ** $p < 0,01$, *** $p < 0,001$

Cuadro 4: Efecto Agregado Incluyendo Variables de Pre-Tratamiento

	G2007	G2008	G2009	G2010	G2011	G2012	G2013	G2014	G2015	G2016	G2017	G2018	G2019	G2020	G2021	G2022	G2023
Prob(GN=1)	-10.63***	-3.61***	-12.78***	-3.95***	-10.94***	-15.71***	-14.27***	-14.05***	-13.31***	-11.79***	-10.27***	-11.26***	-10.42***	-4.72***	-12.57***	-21.06***	-14.89***
	(0.10)	(0.10)	(0.10)	(0.07)	(0.06)	(0.05)	(0.04)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
ingreso neto total	0.00***	0.00***	0.00***	0.00***	0.00***	0.00***	0.00***	0.00***	0.00***	0.00***	0.00***	0.00***	0.00***	0.00***	0.00***	0.00***	0.00***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
que edad tiene en aumplidos? (en a	0.11***	0.13***	0.11***	0.11***	0.13***	0.11***	0.11***	0.11***	0.10***	0.10***	0.09***	0.07***	0.09***	0.07***	0.07***	0.10***	0.15***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
total de miembros del hogar	2.32***	2.28***	2.77***	2.70***	3.09***	3.00***	3.10***	2.83***	2.39***	2.66***	2.44***	2.62***	2.45***	2.53***	3.35***	4.17***	3.57***
	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Edu	1.07***	0.49***	1.20***	0.82***	0.84***	0.55***	0.26***	-0.35***	0.33***	0.83***	-0.42***	0.65***	0.20***	0.25***	0.01	-0.50***	0.95***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
esix	-0.18***	-0.30***	-1.15***	-0.88***	-1.45***	-1.19***	-1.92***	-1.28***	-0.90***	-1.06***	-0.63***	-0.92***	-0.98***	-1.09***	-1.08***	-0.30***	-0.64***
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)
Urbana	-1.47***	0.00	0.00	0.00	-7.36***	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	(0.37)	(.)	(.)	(.)	(0.38)	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)
propia	2.37***	0.77***	1.36***	1.83***	1.38***	1.80***	1.02***	0.88***	-0.00	0.22***	0.99***	1.08***	0.49***	-0.61***	1.35***	1.39***	0.57***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Constant	16.16***	15.40***	12.70***	12.64***	19.98***	14.82***	15.78***	16.99***	19.11***	17.21***	18.23***	18.73***	18.79***	22.19***	28.45***	33.90***	24.20***
	(0.38)	(0.04)	(0.04)	(0.04)	(0.38)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.05)	(0.05)	(0.05)
Observations	1752188	1898582	1922472	2010588	2061400	2145462	2297064	2347182	2480769	2484221	2518336	2592436	2708828	2787868	2968440	2985394	2996609
R ²	0.206	0.161	0.241	0.227	0.212	0.219	0.212	0.219	0.193	0.219	0.205	0.221	0.214	0.130	0.186	0.318	0.248

Standard errors in parentheses
 * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Cuadro 5: Procedimiento de Horváth y Maldoom- Probit

	P2007	P2008	P2009	P2010	P2011	P2012	P2013	P2014	P2015	P2016	P2017	P2018	P2019	P2020	P2021	P2022	P2023
leti																	
inghog2d	0.00*** (0.00)	0.00*** (0.00)	0.00*** (0.00)	0.00*** (0.00)	0.00*** (0.00)	0.00*** (0.00)	0.00*** (0.00)	0.00*** (0.00)	0.00*** (0.00)	0.00 (0.00)	0.00*** (0.00)	-0.00*** (0.00)	-0.00*** (0.00)	-0.00*** (0.00)	-0.00*** (0.00)	-0.00*** (0.00)	-0.00*** (0.00)
Edad	0.01*** (0.00)	0.01*** (0.00)	0.00*** (0.00)	0.01*** (0.00)	0.01*** (0.00)	0.01*** (0.00)	0.01*** (0.00)	0.01*** (0.00)	0.01*** (0.00)	0.01*** (0.00)	0.01*** (0.00)	0.01*** (0.00)	0.01*** (0.00)	0.01*** (0.00)	0.01*** (0.00)	0.01*** (0.00)	0.01*** (0.00)
mieperho	0.04*** (0.00)	-0.01*** (0.00)	-0.05*** (0.00)	0.01*** (0.00)	-0.02*** (0.00)	0.05*** (0.00)	0.02*** (0.00)	0.01*** (0.00)	0.02*** (0.00)	0.08*** (0.00)	0.05*** (0.00)	0.08*** (0.00)	0.12*** (0.00)	0.10*** (0.00)	0.17*** (0.00)	0.10*** (0.00)	0.10*** (0.00)
Edu	0.14*** (0.01)	0.24*** (0.01)	0.00 (.)	0.00 (.)	0.00 (0.00)	0.18*** (0.00)	0.04*** (0.00)	0.07*** (0.00)	-0.01** (0.00)	0.13*** (0.00)	0.11*** (0.00)	0.14*** (0.00)	0.03*** (0.00)	0.02*** (0.00)	0.05*** (0.00)	0.03*** (0.00)	-0.05*** (0.00)
esix	0.05*** (0.00)	-0.07*** (0.01)	-0.02* (0.01)	-0.26*** (0.01)	-0.20*** (0.00)	-0.11*** (0.00)	-0.06*** (0.00)	0.05*** (0.00)	0.03*** (0.00)	-0.12*** (0.00)	-0.05*** (0.00)	-0.02*** (0.00)	-0.09*** (0.00)	-0.06*** (0.00)	-0.15*** (0.00)	-0.06*** (0.00)	-0.13*** (0.00)
propia	0.44*** (0.01)	0.15*** (0.01)	0.22*** (0.01)	-0.12*** (0.01)	0.20*** (0.00)	0.22*** (0.00)	0.09*** (0.00)	0.15*** (0.00)	0.06*** (0.00)	0.14*** (0.00)	0.12*** (0.00)	0.13*** (0.00)	0.24*** (0.00)	0.18*** (0.00)	0.26*** (0.00)	0.29*** (0.00)	0.31*** (0.00)
Pobre	-0.40*** (0.01)	-0.09*** (0.01)	0.00 (.)	0.00 (.)	0.00 (.)	-0.18*** (0.01)	-0.06*** (0.01)	-0.19*** (0.01)	-0.24*** (0.01)	-0.32*** (0.00)	-0.08*** (0.00)	-0.44*** (0.00)	-0.41*** (0.00)	-0.36*** (0.00)	-0.32*** (0.00)	-0.07*** (0.00)	-0.12*** (0.00)
Constant	-3.51*** (0.02)	-3.13*** (0.02)	-2.79*** (0.02)	-2.86*** (0.01)	-2.42*** (0.01)	-2.72*** (0.01)	-2.59*** (0.01)	-2.46*** (0.01)	-2.21*** (0.01)	-1.99*** (0.01)	-2.07*** (0.01)	-2.02*** (0.00)	-1.70*** (0.00)	-1.75*** (0.00)	-1.82*** (0.00)	-1.22*** (0.00)	-1.33*** (0.00)
Observations	2327	2315	1355	1349	2255	2601	3603	3689	3678	3789	3757	3747	3776	3815	3730	3782	3730

Standard errors in parentheses

* $p < 0,05$, ** $p < 0,01$, *** $p < 0,001$

Cuadro 6: Procedimiento de Horváth y Maldoom- Predicción del Gasto

	G2007	G2008	G2009	G2010	G2011	G2012	G2013	G2014	G2015	G2016	G2017	G2018	G2019	G2020	G2021	G2022	G2023
inghog2d	-0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	-0.00 (0.00)	0.00 (0.00)	0.00* (0.00)	-0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00*** (0.00)	0.00*** (0.00)	0.00* (0.00)	0.00** (0.00)	0.00*** (0.00)	0.00 (0.00)
Edad	-0.54 (0.54)	1.55 (1.49)	-0.11 (0.73)	1.63 (1.72)	-0.31 (0.30)	-0.53* (0.21)	0.03 (0.28)	0.23 (0.25)	-0.42 (0.24)	-0.11 (0.08)	-0.34** (0.12)	-0.15 (0.08)	-0.03 (0.05)	-0.21** (0.07)	-0.05 (0.06)	-0.12* (0.05)	-0.25*** (0.06)
Sexo	-4.96 (10.03)	18.63 (9.76)	9.65 (19.72)	14.07 (15.08)	-7.47 (11.87)	-1.65 (3.65)	0.74 (2.90)	-3.49 (2.69)	-1.11 (2.16)	-0.49 (2.12)	4.07* (1.82)	2.78* (1.15)	2.23 (1.25)	0.75 (1.45)	-0.78 (1.36)	-0.28 (1.09)	-0.14 (1.07)
Edu	12.38 (18.34)	-6.06 (31.72)	0.00 (.)	0.00 (.)	1.68 (8.91)	-10.67 (5.52)	3.11 (2.62)	-2.89 (2.56)	1.26 (2.01)	-2.02 (1.76)	-7.66*** (2.13)	-5.21*** (1.44)	-0.41 (1.21)	0.38 (1.37)	-1.40 (1.48)	-1.37 (1.18)	2.84* (1.17)
esix	9.53 (6.67)	10.51 (9.65)	5.16 (19.61)	-5.38 (30.77)	-0.30 (7.77)	-0.73 (2.26)	-0.48 (1.74)	1.32 (2.06)	-1.22 (1.64)	2.45* (1.10)	1.24 (1.09)	-0.09 (0.89)	0.89 (0.98)	-0.10 (1.12)	0.60 (1.58)	2.39* (1.10)	3.21*** (0.83)
RIM1	-73.81 (383.47)	-809.53 (2058.46)	-179.54 (191.24)	-703.95 (675.57)	9.14 (125.74)	320.81** (114.83)	32.10 (90.25)	-65.17 (84.21)	128.67* (63.10)	67.45*** (18.98)	90.31*** (23.80)	39.41** (12.68)	34.76*** (7.13)	57.32*** (8.51)	27.50*** (5.17)	73.19*** (8.84)	72.75*** (6.92)
Pobre	3.81 (11.55)	-41.69 (24.97)	0.00 (.)	0.00 (.)	0.00 (.)	4.05 (3.03)	3.85 (3.40)	-2.86 (4.16)	3.06 (4.10)	10.75** (3.40)	2.49 (2.63)	8.49** (2.94)	5.54** (1.81)	4.67** (1.73)	3.51 (1.89)	-2.49* (1.26)	-0.20 (1.10)
propia	46.05 (23.12)	21.90 (22.16)	-6.81 (16.07)	-6.87 (13.91)	4.41 (8.22)	-15.36** (5.66)	-2.80 (4.00)	3.18 (3.81)	-0.76 (2.23)	-4.45* (1.89)	-3.37 (1.99)	-1.17 (1.40)	-4.59** (1.43)	-5.57*** (1.51)	-1.64 (1.52)	-9.88*** (1.65)	-10.92*** (1.44)
EC	2.06 (4.75)	11.56 (5.35)	7.16 (7.66)	-3.35 (3.95)	-2.26 (3.03)	2.44 (1.34)	0.17 (0.80)	-1.31 (0.70)	-1.53** (0.52)	0.10 (0.73)	0.62 (0.64)	-0.19 (0.36)	0.24 (0.41)	0.07 (0.44)	-0.66 (0.40)	-0.52 (0.31)	-0.74* (0.35)
nbi1	0.00 (.)	0.00 (.)	0.00 (.)	0.00 (.)	0.00 (.)	0.00 (.)	4.38 (3.54)	0.00 (.)	0.00 (.)	-4.26* (1.97)	0.00 (.)	7.86 (9.40)	-0.43 (2.85)	-1.41 (4.90)	2.07 (6.60)	23.12 (12.14)	-7.69*** (2.28)
Constant	7.80 (33.22)	-93.23 (80.29)	10.22 (17.75)	-24.19 (64.09)	43.70* (20.77)	38.92** (12.13)	14.52 (8.92)	16.72 (8.97)	29.47*** (7.24)	13.28** (4.26)	20.07*** (4.97)	17.05*** (2.69)	9.29** (2.96)	16.60*** (3.69)	23.21*** (3.81)	2.12 (4.13)	5.52 (3.30)
Observations	17	15	11	28	43	75	183	224	298	507	532	662	816	920	982	1298	1436

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Cuadro 7: Propensity Score Matching

	Año	T	C	NNM	t	T	C	RM (0.2)	t	RM (0.1)	t	RM (0.05)	t	KM
r1	2007.0	17.0	17.0	-6.9	-1.9	17.0	17.0	-6.9	-1.9	-6.9	-1.9	-7.0	-2.0	-7.0
r2	2008.0	15.0	15.0	1.1	0.2	15.0	15.0	1.1	0.2	1.1	0.2	1.1	0.2	1.1
r3	2009.0	11.0	11.0	-7.4	-1.7	11.0	11.0	-7.4	-1.7	-7.4	-1.7	-7.4	-1.7	-7.4
r4	2010.0	28.0	28.0	1.3	0.2	28.0	28.0	1.2	0.2	1.2	0.2	1.2	0.2	1.2
r5	2011.0	43.0	43.0	-9.3	-2.3	43.0	43.0	-9.3	-2.3	-9.3	-2.3	-9.3	-2.3	-9.3
r6	2012.0	75.0	73.0	-13.5	-7.2	75.0	75.0	-13.9	-7.3	-13.9	-7.3	-13.7	-7.2	-13.8
r7	2013.0	183.0	179.0	-11.2	-7.5	183.0	183.0	-11.3	-7.5	-11.3	-7.5	-11.5	-7.6	-11.4
r8	2014.0	224.0	215.0	-13.0	-9.5	224.0	224.0	-12.7	-9.5	-12.7	-9.5	-12.8	-9.5	-12.8
r9	2015.0	298.0	294.0	-12.7	-12.6	298.0	298.0	-12.8	-12.8	-12.8	-12.8	-12.8	-12.5	-12.8
r10	2016.0	508.0	476.0	-12.3	-13.4	508.0	508.0	-12.2	-13.8	-12.0	-13.4	-12.0	-13.2	-12.1
r11	2017.0	532.0	509.0	-10.1	-12.2	532.0	532.0	-10.2	-12.2	-10.1	-12.0	-10.1	-11.9	-10.3
r12	2018.0	662.0	634.0	-10.7	-14.6	662.0	662.0	-10.9	-15.1	-10.5	-14.5	-10.4	-14.2	-10.7
r13	2019.0	816.0	715.0	-10.7	-15.0	816.0	816.0	-10.5	-15.9	-10.4	-15.4	-10.3	-15.1	-10.6
r14	2020.0	920.0	804.0	-4.5	-6.2	920.0	920.0	-4.5	-6.5	-4.6	-6.7	-4.9	-7.0	-4.5
r15	2021.0	982.0	885.0	-12.9	-16.2	982.0	982.0	-12.6	-16.8	-12.4	-16.2	-12.2	-15.9	-12.8
r16	2022.0	1298.0	1104.0	-21.7	-25.8	1298.0	1298.0	-22.0	-29.5	-22.1	-29.1	-22.3	-29.0	-21.9
r17	2023.0	1436.0	1189.0	-14.2	-18.0	1436.0	1436.0	-14.9	-23.0	-15.0	-22.9	-15.2	-23.0	-14.8

T=Tratamiento,C= Control

NNM = Nearest neighbor matching, t = t de student, RM =Radius matching, KM=Kernel matching