

Relación entre el Sistema de Salud Público y la Informalidad en Colombia

Sebastián Vásquez Orellana

Trabajo de grado

Asesor: Santiago Bohorquez Correa

Universidad EAFIT

Escuela de Finanzas, Economía y Gobierno

Pregrado en Economía

Medellín, Colombia

2023

Resumen

El trabajo se propone investigar la relación existente entre el sector informal y el sistema de salud en Colombia, dada la vinculación entre la protección social con el estatus laboral del individuo, que genera una serie de incentivos sobre los actores del mercado laboral. Específicamente, se plantea la hipótesis de que pertenecer al régimen subsidiado incrementa la probabilidad de pertenecer al sector informal del individuo. Para probarlo, se emplean modelos lineales de probabilidad, logit y probit junto con un modelo LASSO para la selección de las variables. Se encontró que pertenecer al RS está asociado con una probabilidad mayor en 13 pp a ser informal. Sin embargo, para reflejar de manera adecuada la relación de doble causalidad entre las variables de interés, se propone un modelo probit bivariado. Los resultados permiten concluir la existencia de una relación significativa entre la informalidad y el sistema de salud.

Palabras clave: Informalidad, régimen de salud, mercado laboral, incentivos.

JEL code: E26, J32

Abstract: This paper aims to investigate the relationship between the informal sector and the health system in Colombia, given the link between social protection and the labor status of the individual, which generates a series of incentives on the actors in the labor market. Specifically, it is hypothesized that belonging to the subsidized system increases the probability of belonging to the informal sector. To test this, linear probability, logit and probit models are used along with a LASSO model for the selection of variables. It was found that belonging to the subsidized regime is associated with a 13 pp higher probability of being informal. However, to adequately reflect the dual causal relationship between the variables of interest, a bivariate probit model is proposed. The results allow us to conclude the existence of a significant relationship between informality and the health system.

Contenido

Introducción	4
Contextualización	5
Informalidad en Colombia	5
Sistema de Salud Público en Colombia	9
Relación entre la informalidad y el sistema de salud.....	11
Datos empleados	14
Metodología	16
LASSO (Least Absolute Shrinkage and Selection Operator)	16
Probit bivariado.....	17
Resultados	19
Conclusión	23
Referencias.....	25
Apéndice A	29
Apéndice B.....	31
Apéndice C.....	33

Introducción

La informalidad es uno de los problemas más latentes del mercado laboral, de él se derivan múltiples repercusiones que impactan la calidad de vida de los trabajadores y, de manera más amplia, la economía. Históricamente, los niveles de informalidad en Colombia han sido altos. Para el último periodo que se tienen datos, noviembre 2022-enero 2023, la proporción de ocupados informales fue de 42,3% para las 13 ciudades y áreas metropolitanas (Departamento Administrativo Nacional de Estadística, 2023).

Por una parte, preocupa la falta de protección contra enfermedad, accidentes, invalidez o muerte del trabajador (Camacho et al., 2013; Perry et al., 2007) al igual que la dificultad que la persona adquiera experiencia y pueda mejorar sus ingresos a lo largo de su ciclo de vida (Alvarado, Meléndez, et al., 2021; Santa María et al., 2009), perpetuando la condición de pobreza y falta de educación que impide la salida de la informalidad. Por otra, se observa que esta frena el crecimiento económico y desarrollo, mediante la baja productividad, característica de las empresas que se desempeñan dentro del sector (Mondragón-Vélez et al., 2010; Perry et al., 2007; Santa María et al., 2009); y el menor recaudo tributario que hace que el Gobierno disponga de menores ingresos fiscales. (Perry et al., 2007; Vásquez et al., 2022).

Estas problemáticas están estrechamente ligadas ya que, el deficiente desempeño económico afecta la prosperidad de los trabajadores. Por tal motivo, resulta imperante analizar los factores relacionados con este fenómeno, con el fin de comprenderlo a profundidad y generar propuestas que permitan avanzar en las dimensiones de bienestar social y crecimiento económico. Para el caso colombiano, el diseño actual del sistema de salud, mediante sus dos regímenes puede estar relacionado con el estatus laboral del individuo. Por tal motivo, se introduce el Sistema General de Seguridad Social en Salud (SGSS) dentro del análisis propuesto.

La falencia se encuentra en que los altos costos no salariales¹ (CNS), en los que está incluido el aporte a la salud, hacen que los empleadores deban disminuir su demanda de trabajo, debido a la existencia del salario mínimo que imposibilita una reducción vía precios (Kugler & Kugler, 2003; Mondragón-Vélez et al., 2010; Santa María et al., 2009). Adicionalmente, la existencia del Régimen Subsidiado (RS) puede suponer un incentivo a la informalidad porque

¹ Se consideran costos no salariales los pagos realizados por concepto de cesantías, pensiones, salud, parafiscales y primas (Santa María et al., 2009).

ofrece los mismos beneficios que el Régimen Contributivo (RC) de manera gratuita, haciendo que las personas elijan la informalidad (cuenta-propia o trabajo en empresas pequeñas) (Mora & Muro, 2014; Saavedra-Caballero & Ospina Londoño, 2018), dado que dentro de su función de utilidad los costos de pertenecer al RC son mayores que los beneficios que reciben.

El salario mínimo y los altos CNS representan unas rigideces de mercado que han hecho que en la actualidad el mercado laboral colombiano esté profundamente segmentado (Alvarado, Meléndez, et al., 2021; Mondragón-Vélez et al., 2010). Bajo este contexto se presentan efectos de expulsión y de salida voluntaria del mercado formal hacia el mercado informal, en donde los más afectados son aquellas personas con bajos ingresos, poca educación, ser mano de obra no especializada y cuyo salario oscila alrededor del salario mínimo -es decir, casi la mitad de la población ocupada- (Alvarado, Álvarez, et al., 2021; Perry et al., 2007; Santa María et al., 2009).

Por tal motivo, se analiza la relación entre la probabilidad de pertenecer al régimen subsidiado y la de ser informal. Adicionalmente, se pretende probar que las(os) cónyuges de los(as) jefes(as) de hogar tienen una mayor probabilidad de ser informales. Para esto, se emplearon datos de la Gran Encuesta Integrada de Hogares (GEIH) del DANE a nivel nacional entre 2015 y 2019. Haciendo una selección de variables mediante un modelo LASSO que permita ajustar el poder predictivo de los modelos probit, logit y de probabilidad lineal empleados. Finalmente, se busca una mayor exactitud a través de un modelo probit bivariado ya que captura de manera más adecuada la relación entre las variables de interés.

El documento se divide en siete partes, siendo la primera esta introducción. A continuación, se realiza una contextualización de la situación actual de informalidad en Colombia junto con una breve explicación del sistema de salud; la tercera parte presenta la revisión de literatura acerca de la relación entre las dos variables; la cuarta muestra los datos empleados junto con sus principales estadísticas descriptivas; la quinta discute la metodología empleada; la sexta muestra y analiza los resultados obtenidos; y finalmente se concluye en la séptima parte.

Contextualización

Informalidad en Colombia

Para hablar del fenómeno de la informalidad en Colombia, es necesario hacer algunas aclaraciones sobre las diferentes definiciones que se tienen para describirlo. En primer lugar, se

debe comprender que la informalidad es un atributo que se le asigna a la(s) actividad(es) que desempeña un individuo, mas no al individuo mismo (de Soto, 1987). Esto implica que es un estatus transitorio por el que puede atravesar el individuo durante varias etapas de su vida. De acuerdo con la Organización Internacional del Trabajo (OIT, s.f.) , el empleo informal:

Incluye todo el trabajo remunerado (es decir, tanto el empleo por cuenta propia como el asalariado) que no está registrado, regulado o protegido por los marcos jurídicos o normativos existentes, así como el trabajo no remunerado realizado en una empresa generadora de ingresos. Los trabajadores informales carecen de contratos de trabajo seguros, prestaciones laborales, protección social o representación de los trabajadores. (p. 1)

Sin embargo, a veces cuantificar la población con un empleo informal puede ser complicado dada la complejidad del fenómeno. Bernal (2009), identifica 27 formas diferentes para cuantificar el nivel de informalidad, clasificadas en tres grupos: acceso a la seguridad social, existencia de un contrato formal y tamaño de la empresa. Esto es similar a la propuesta de Gasparini & Tornarolli (2007), que sugieren una definición “productiva”, que captura los trabajadores de sectores marginales, de baja productividad, con una alta demanda por mano de obra; y una definición “legalista” que se refiere a aquellos sin protección laboral u otros beneficios de la seguridad social.

Para el caso de Colombia, el marco de la GEIH 2005 se basa en la definición de tamaño de la empresa y ocupación del individuo, tomando en consideración la definición de “sector informal” de la OIT y las recomendaciones del Grupo Delhi (Departamento Administrativo Nacional de Estadística, 2022). Como resultado, se consideraba como ocupado en el sector informal a quien cumpliera una de las siguientes condiciones²:

- i) Los empleados particulares y los obreros que laboran en establecimientos, negocios o empresas que ocupen hasta cinco personas en todas sus agencias y sucursales, incluyendo al patrono y/o socio;
- ii) Los trabajadores familiares sin remuneración;
- iii) Los trabajadores sin remuneración en empresas o negocios de otros hogares;

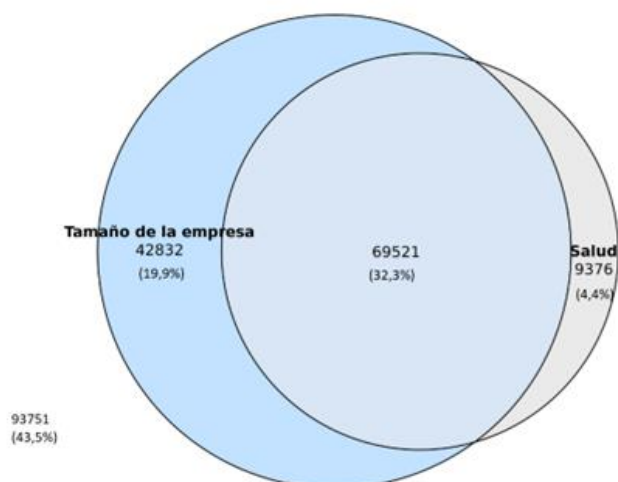
² Esta definición aplica únicamente para los dominios urbanos

- iv) Los empleados domésticos;
- v) Los jornaleros o peones;
- vi) Los trabajadores por cuenta propia que laboran en establecimientos hasta cinco personas, excepto los independientes profesionales;
- vii) Los patronos o empleadores en empresas de cinco trabajadores o menos;
- viii) Esta medida excluye los obreros o empleados del gobierno.

Para el nuevo marco 2018 de la GEIH, se considera como ocupado dentro del sector informal a aquel trabajador que se desempeñe en una empresa sin registro mercantil de la Cámara de Comercio y a aquel asalariado o empleado domestico que no cuente con cotización de salud ni pensión; así como trabajadores familiares sin remuneración, los trabajadores por cuenta propia y patronos o empleadores que hayan quedado clasificados en el sector informal (Departamento Administrativo Nacional de Estadística, 2022). Este marco comenzó a ser efectivo a partir de 2021.

Como el propósito de este trabajo es investigar el efecto que tiene la afiliación al RS sobre la probabilidad de estar ocupado en el sector informal de un individuo, se emplea la definición del marco 2005. Bernal (2009) afirma que la definición de informalidad por tamaño de la empresa tiene una correlación fuerte con aquella de afiliación a pensión y salud. Sin embargo, Mondragón-Vélez et al. (2010) encuentran diferencias relevantes entre definiciones de tamaño de la empresa y afiliación a salud para el caso de Colombia en 2006, aproximadamente 23% de los trabajadores eran formales bajo una definición pero informales bajo la otra.

Siguiendo el análisis de Mondragón-Vélez et al. (2010), se observa la concurrencia de ambas definiciones para los datos empleados. Se encuentra que, bajo la definición de tamaño de la empresa, entre 2015 y 2019, un promedio del 47,15% estuvo ocupado en la informalidad. Mientras que, bajo la definición de afiliación a salud, tan solo un 30,42% se consideran informales. Se debe aclarar que los datos empleados corresponden a las 13 ciudades principales y áreas metropolitanas. Adicionalmente, el 32,3% de las observaciones se desempeñan en el sector informal bajo ambas definiciones, y la definición de tamaño de la empresa captura a la mayoría de los trabajadores considerados informales bajo la definición de afiliación a salud.

Figura 1*Convergencia de las definiciones de informalidad*

Nota. Se observa la convergencia entre las definiciones de informalidad para todas las observaciones entre 2015 y 2019. Elaboración propia con datos de la *Gran Encuesta Integrada de Hogares* (GEIH). DANE (2021).

Finalmente, hay que mencionar que la informalidad puede ser un mecanismo complementario o sustituto del mercado laboral formal dependiendo de la manera en que los individuos se adhieran a este (Santa María et al., 2009). Por una parte, un individuo puede pertenecer al sector informal como resultado de un proceso de exclusión, generado, entre múltiples factores, por las rigideces del mercado formal – la existencia de un salario mínimo y altos CNS-. En este sentido se tendría un mercado laboral segmentado, con un sector informal que sirve como sustituto del sector formal y que se mueve en dirección contraria al ciclo económico (Mondragón-Vélez et al., 2010).

Por otra, el individuo puede haber elegido salir del mercado formal, dado que bajo su análisis representa unos costos mayores que los beneficios que puede recibir al estar bajo una relación de subordinación (Perry et al., 2007). En este caso, el sector informal sería un complemento al formal y sería procíclico. Para el caso de Colombia, los resultados indican un mercado laboral segmentado, donde el sector informal recoge en su mayoría de trabajadores

expulsados del sector formal. Mondragón-Vélez et al. (2010) afirma que el sector informal es contra cíclico. De manera similar, Bernal (2009) encuentra que alrededor de la mitad de los trabajadores informales aceptarían un trabajo formal por un salario igual o menor al que tienen actualmente. Por último, en la Encuesta Social Longitudinal de Fedesarrollo (2009) se observa una mayor disposición a perder la afiliación al RS con el fin de obtener un empleo formal.

Sistema de Salud Público en Colombia

Mediante la Ley 100 de 1993 se creó el Sistema General de Seguridad Social en Salud (SGSSS) que funciona actualmente en el país. Para 2022, el Ministerio de Salud y Protección Social (Minsalud) reportó una cobertura del 99,6% de la población. La afiliación a este sistema es de carácter obligatorio y se realiza a través de las Entidades Promotoras de Salud (EPS), quienes reciben las cotizaciones y ofrecen el Plan Obligatorio de Salud (POS) a través de las Instituciones Prestadoras de Servicios (IPS) (Guerrero et al., 2011). Con la intención de garantizar la cobertura universal en salud, se conformaron dos regímenes: el contributivo y el subsidiado.

El régimen contributivo (RC) está conformado por todos los trabajadores, pensionados y trabajadores independientes que devenguen más de 1 SMMLV -y sus beneficiarios- que de forma obligatoria deben afiliarse y se sostiene con los aportes que estos mismos realicen a la EPS a la cual decidan afiliarse libremente. En la actualidad, luego de la Ley 1122 de 2007, este aporte es del 12,5%. De esta tasa, el empleado debe cubrir el 4% mientras que el empleador el 8,5% restante -los trabajadores independientes deben cubrir la totalidad del monto-. De este monto, 1,5 puntos porcentuales se destinan al Fondo de Solidaridad y Garantía -FOSYGA- a cargo del pago de las tutelas y recobros; 0,5 van a la Subcuenta de la Promoción de la Salud y 1pp van a la Subcuenta de Solidaridad, para contribuir a la financiación del Régimen Subsidiado de salud (Alvarado, Meléndez, et al., 2021). De acuerdo con el Ministerio de Salud (2022), 24'399.839 de personas pertenecen al RC.

Por otro lado, está el régimen subsidiado (RS) al cual pertenecen los grupos poblacionales sin capacidad de pago, este es financiado en parte con aportes de los afiliados al RC y por el Estado. La selección de los beneficiarios se da a través del Sistema de Identificación de Potenciales Beneficiarios de Programas Sociales (SISBEN), donde ingresan a los subsidios totales quienes se encuentren en los grupos A1-B7 y C1-C18 de la clasificación. A este régimen pertenecen

24'745.934 de personas (Ministerio de Salud, 2022). El SISBEN IV clasifica los hogares en cuatro categorías de acuerdo con sus condiciones de vida e ingresos teniendo en cuenta variables de vivienda y servicios públicos, educación, salud, ocupación y antecedentes sociodemográficos (Departamento Nacional de Planeación (DNP), s.f.).

Dentro del sistema de salud público existen tanto cuotas moderadoras como copagos, las primeras son aplicables únicamente a los cotizantes y a sus beneficiarios del RC, mientras que los copagos se aplican a los beneficiarios del RC y a los afiliados del RS. Los montos de estos desembolsos son tabulados de acuerdo con el servicio requerido y el grupo de ingresos del individuo (Guerrero et al., 2011) y su objetivo es “racionalizar el uso de servicios del sistema” y “complementar la financiación del Plan Obligatorio de Salud (POS)” (Ley 100 de 1993). Este gasto de bolsillo ha venido disminuyendo desde el 2010, y para 2017 representó el 16,3% del gasto total en salud, siendo de los países con menor gasto de bolsillo (Organización para la Cooperación y el Desarrollo Económicos & Banco Mundial, 2020).

El gasto público constituye la mayor parte del gasto total en salud (75%), (Minsalud, 2022), respecto al PIB, se ubicó en promedio en 5,5% para el periodo 2015-2020. El incremento del gasto público corresponde con el incremento de la cobertura del RS y con la evolución de la Unidad de pago por Capitación (UPC) del mismo luego de la sentencia T-760 de 2008. Dicha sentencia decreta la igualación de los beneficios en salud entre los dos regímenes, con la única diferencia de que en el RS no se reconocen el subsidio en dinero en caso de incapacidad, ni la licencia de maternidad (o paternidad). Como resultado, el valor anual que se reconoce por cada uno de los afiliados (UPC) es casi el mismo, sin importar a cuál régimen pertenezca (Alvarado, Meléndez, et al., 2021).

La sostenibilidad del sistema se dificulta dado el incentivo a mantener las condiciones para hacer parte del RS mediante la informalidad, como se verá más adelante. De acuerdo con García (2015, citado en Consejo Privado de Competitividad, 2017), se dejan de recaudar \$4 billones por concepto de informales con capacidad de pago que deciden pertenecer al RS. En el [Apéndice A](#) se encuentran gráficos relevantes para esta caracterización del sistema de salud colombiano.

Relación entre la informalidad y el sistema de salud

Ahora, la problemática estructural que se presenta entre estas dos variables es la asociación de la protección social al estatus laboral del trabajador, e indirectamente a sus ingresos (Alvarado, Álvarez, et al., 2021). Esto hace que las empresas deban cubrir riesgos ajenos a ellas, como los de enfermedad, invalidez, longevidad y muerte. Como se mencionó en la introducción, el incremento de los CNS como consecuencia de este marco regulatorio hace que los empleadores deban modificar sus decisiones en el mercado laboral y, como existe un salario mínimo que dificulta la transmisión de estos costos vía precios, se ven obligados a reducir su demanda. De acuerdo con Santa María et al., (2009), la Ley 100 de 1993 incrementó los CNS a 58,1% del salario mensual.

La reducción del número de personas en el sector formal como consecuencia de unos altos CNS, es evidencia de un mercado laboral segmentando. En donde principalmente los trabajadores que ganan un salario mínimo son expulsados del sector formal hacia el informal, estos son individuos que, como ya se mencionó, se caracterizan por tener poca educación y ser mano de obra no especializada. El hecho de que el salario mínimo sea un determinante para el aporte a la salud hace que el sistema sea altamente excluyente y no tome en cuenta la situación de casi la mitad de la población -que gana menos de dicho salario- (Alvarado, Álvarez, et al., 2021; Santa María et al., 2009).

Adicional a esto, la existencia de un sistema de seguridad gratuito influye sobre las decisiones de empleo de los individuos, si este tiene una valoración baja de la utilidad que le representa tener un mejor acceso a la salud, preferirá ocuparse en el sector informal (Bernal, 2009; Camacho et al., 2013; Levy, 2008; Santa María et al., 2011). Camacho et al., 2013, encuentran que jefes de hogares con adultos mayores y niños tienen una mayor valoración del acceso a la salud y por lo tanto tienden a preferir el RC; contrario a los jóvenes, sin dependientes, que tienden a preferir el RS. Sin embargo, luego de la sentencia T-760 de 2008 que ordena la igualación de los beneficios en salud entre regímenes, no hay ningún incentivo para permanecer en el RC.

El RS también impacta la decisión de las empresas, dado que sus trabajadores pueden mantener los beneficios en salud si esta se mantiene lo suficientemente pequeña, evadiendo la ley e impactando la productividad de la economía (Alvarado, Meléndez, et al., 2021). En el caso de los trabajadores independientes que ganan menos del salario mínimo, no se presenta ningún incentivo a contribuir al RC, dado que las cotizaciones se realizan con base a este salario mínimo,

por lo que tendrían que pagar una proporción mayor de sus ingresos (Alvarado, Álvarez, et al., 2021). Mientras que los que ganan más del SMMLV se ven incentivados a simular sus respuestas del SISBEN y mantenerse dentro del RS y el paquete de ayudas adicionales que otorga el Estado (Camacho et al., 2013). De acuerdo con Saavedra-Caballero & Ospina Londoño (2018), recibir el subsidio de Familias en acción, incrementó en 7,13 pp la probabilidad de pertenecer al régimen subsidiado, y por consiguiente, de ser informal. Para el caso de los cónyuges del jefe de hogar, no se observa ningún incentivo a pertenecer al sector formal, dado que de hacerlo, el hogar percibiría una doble cotización en salud (Mondragón-Vélez et al., 2010).

Como se mencionó antes, el salario mínimo es un agravante del panorama, dado que parece afectar a más personas de las que beneficia (Alvarado, Álvarez, et al., 2021). Esto, debido a que es un salario relativamente alto en relación con el salario medio de los trabajadores (Figura 2), y en un contexto de altos CNS, representa una barrera de acceso al sector formal, generando altas tasas de informalidad y desempleo. En este sentido resulta poco eficiente como instrumento redistributivo (Kugler & Kugler, 2003; Mondragón-Vélez et al., 2010; Santa María et al., 2009). Adicionalmente, se emplea para indexar los incrementos salariales de toda economía, esto impacta negativamente a los individuos cuyos salarios oscilan alrededor de este, ya que genera inflexibilidades salariales (Maloney & Nuñez, 2001).

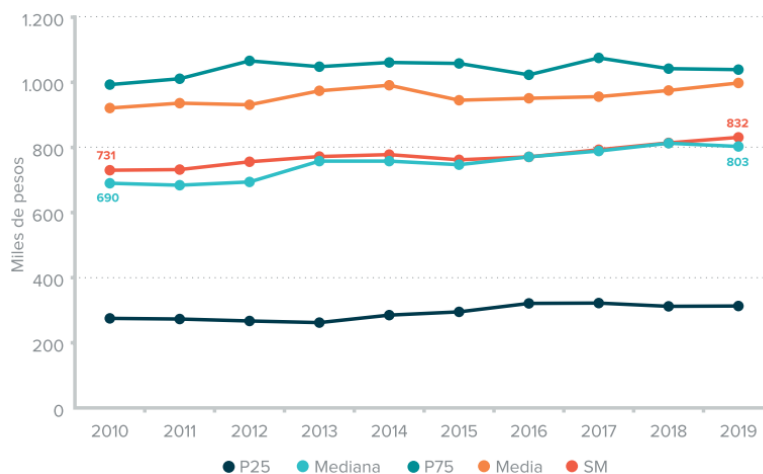
La dificultad de transición entre el sector formal e informal también es una rigidez que hace que los individuos prefieran mantenerse en el último. Dado que de ingresar al sector formal y cambiarse al RC, tendrían grandes dificultades para reintegrarse al RS en caso de que pierdan su capacidad de pago -pierdan su empleo- (Bernal, 2009; Mondragón-Vélez et al., 2010). Solo aquellos que pertenezcan a los grupos A1-B7 y C1-C18 del SISBEN tienen la posibilidad de pasar directamente al RS, de lo contrario se deberá solicitar la aplicación de la encuesta (Minsalud, 2014). El estudio de Canavire-Bacarreza et al. (2017) concluye que es más probable que un trabajador informal permanezca en este segmento del mercado laboral, y que trabajadores que previamente eran formales entren a la informalidad, a que un trabajador informal pueda salir del sector.

La literatura previa para el caso colombiano se ha centrado principalmente en el efecto de los CNS sobre variables relevantes del mercado laboral como el desempleo, la informalidad y el nivel de salarios. Santa María et al. (2009) encuentran una reducción del 3,9% en la probabilidad de ser

formal dado un incremento del 10% en los CNS. Por su parte, Kugler & Kugler (2003) indican que esta reducción está entre el 4 y 5%, y Mondragón-Vélez et al. (2010) sugieren un incremento entre el 5 y 8% sobre la probabilidad de ser informal.

Figura 2

Salario mínimo e ingreso laboral observado



Tomado de *Mercados laborales fragmentados y el sistema de protección social en Colombia* (p.32), por Alvarado, Meléndez et al., 2021, UNDP LAC Working Paper Series.

Son menos los estudios que hacen lo correspondiente únicamente considerando la afiliación a la salud. Santa María et al. (2011) sugieren que el RS disminuye la probabilidad de tener un empleo formal entre 2 y 9 puntos porcentuales para la población en pobreza extrema, esto posiblemente debido a que existe autoselección para pertenecer a este (Ramírez Hassan et al., 2013). Por su parte, Gaviria et al. (2006), encuentran que la existencia de dicho régimen disminuye la participación laboral en un 24% y Camacho et al. (2013) afirman que hubo un incremento del 4% en la informalidad luego de la implementación de la Ley 100.

En últimas, la coexistencia de estos dos sistemas supone un fuerte incentivo a la informalidad (Alvarado, Meléndez, et al., 2021; Camacho et al., 2013; Vásquez et al., 2022). En el largo plazo, un creciente RS que demanda más recursos, aumentará a su vez los porcentajes de cotización para el RC, generando un círculo vicioso en el que más personas se verán expulsadas del sector formal y empleadas en el informal (Santa María et al., 2011). Esto no es el resultado de una sola reforma, sino de una serie de reformas individuales, que no han tenido la suficiente

coherencia para generar un cambio efectivo (Alvarado, Álvarez, et al., 2021). En los últimos treinta años, se han observado 93 reformas relevantes, lo que en promedio sería una cada cuatro meses.

Diferentes propuestas se han presentado para hacer frente a esta problemática, la Misión de Empleo de 2021 contempla la universalidad en la protección social independientemente del nivel de ingresos de la persona, separando así los programas que son ajenos a la relación de dependencia y subordinación entre empresas y trabajadores. Esto, sumado con una tributación que permita la solidaridad de los que tienen mayores ingresos hacia los que más lo necesiten y unos subsidios que incentiven el trabajo formal. Por su parte, la OCDE (2022) propone la economía social, particularmente las cooperativas, como una herramienta para combatir la informalidad. Kugler & Kugler (2003), afirman que la reducción en los aportes a rubros que no benefician directamente al trabajador -como los parafiscales- ayudaría a impulsar la demanda de mano de obra poco calificada.

Finalmente, se debe aclarar que la descripción realizada está enfocada únicamente en un aspecto del sistema de salud, el incentivo a la informalidad, que representa un costo de la estructuración actual. Camacho et al. (2013) encuentran una reducción del 3,8% en el PIB como consecuencia del incremento en la informalidad luego de la implementación de la Ley 100. Mientras que García (2015, citado en Consejo Privado de Competitividad, 2017), afirma que se dejan de recaudar \$4 billones por concepto de informales con capacidad de pago que deciden pertenecer al RS.

Para una evaluación completa del sistema y recomendaciones de política pública es necesario comparar los beneficios con los costos que se generan. Por ejemplo, Gaviria et al. (2006) encuentran que la afiliación al RS mejoró el estado de salud reportado por el individuo, asociado con la mayor facilidad de acceder a consultas médicas preventivas. También se observó una disminución en la frecuencia de las hospitalizaciones

Datos empleados y estadísticas descriptivas

Como ya se mencionó, se tomaron datos mensuales de la GEIH entre 2015 y 2019 y se hizo un *pooling* de estos. La razón de elegir el periodo seleccionado radica en que, para entonces aún se empleaba la definición de tamaño de la empresa y, porque de enero a abril del 2020 el DANE excluyó de la encuesta variables relevantes para la caracterización de la informalidad, debido a la

emergencia sanitaria. Se tomó una muestra del 25% de los datos (aprox. 205 mil observaciones), se construyó la variable de informalidad mediante la definición del DANE que será la variable dependiente y se seleccionaron las variables independientes a través de la revisión de literatura y del método LASSO. La principal relación por observar es entre la afiliación al régimen de salud y la probabilidad de ser informal. En el [Apéndice B](#) se encuentra la descripción de las 20 variables tenidas en cuenta. A continuación, se presentan las estadísticas descriptivas de las 6 variables seleccionadas.

Tabla 1

Estadísticas descriptivas

Variable	Total muestra				Informal=0			
	Promedio	Desv. Es	Min	Max	Promedio	Desv. Es	Min	Max
<i>contrato</i>	0,4058497	0,4910569	0	1	0,8012806	0,3990384	0	1
<i>meses</i>	72,23726	98,83967	0	876	58,99479	81,31609	0	864
<i>horastrab</i>	43,43829	18,10632	0	130	45,08667	14,91451	0	130
<i>regimen</i>	0,3166934	0,4651879	0	1	0,0705257	0,2560322	0	1
<i>edad</i>	39,33926	13,60017	15	100	36,29323	11,60786	15	98
Informal=1								
Variable	Promedio	Desv. Es	Min	Max				
<i>contrato</i>	0,0512408	0,2204896	0	1				
<i>meses</i>	84,11265	110,915	0	876				
<i>horastrab</i>	41,96009	20,43859	0	130				
<i>regimen</i>	0,5374481	0,498598	0	1				
<i>edad</i>	42,07083	14,63631	15	100				

Elaboración propia con datos de la GEIH.

Tabla 2

Porcentaje categorías variable jefe(a) del hogar

Jefe(a)	Parentesco jefe(a) del hogar		
	Total muestra	Informal = 0	Informal=1
Jefe(a)	44,27%	42,73%	45,65%

Cónyuge	19,43%	17,64%	21,03%
Hijo(a)	23,65%	27,90%	19,84%
Nieto(a)	1,68%	1,86%	1,51%
Otro pariente	8,16%	7,65%	8,63%
Empleado(a)	0,73%	0,01%	1,38%
Pensionista	0,09%	0,12%	0,06%
Trabajador	0,06%	0,02%	0,09%
Otro no pariente	1,94%	2,07%	1,82%

Elaboración propia con datos de la GEIH.

Metodología

LASSO (Least Absolute Shrinkage and Selection Operator)

Es un método de estimación de modelos lineales propuesto Tibshirani (1996). Permite la contracción de algunos coeficientes a cero cuando se tiene un alto número de variables con el propósito de reducir la especificidad del modelo y lograr su regularización para una predicción más precisa en datos nuevos. Es particularmente útil para modelos que muestran multicolinealidad o para selección de variables, generando modelos fáciles de interpretar (James et al., 2013)

LASSO emplea una penalización L1 que es igual al valor absoluto de la magnitud del coeficiente, haciendo que algunos coeficientes sean iguales a cero. Matemáticamente, lo que LASSO hace es minimizar la suma residual de cuadrados sujeto a una restricción λ que limita el tamaño de la suma del valor absoluto de todos los coeficientes:

$$\sum_{i=1}^n (y_i - \sum_j x_{ij} \beta_j)^2 + \lambda \sum_{j=1}^p |\beta_j|$$

El *trade-off* entre sesgo-varianza se equilibra mediante el parámetro λ . Si $\lambda = 0$, no se elimina ninguna variable y el estimador será igual al de una regresión lineal, mientras que entre más crezca λ , se irán reduciendo más coeficientes a cero y el sesgo del modelo incrementará. Finalmente, a medida que λ disminuya, la varianza incrementará.

En el ejercicio realizado, se seleccionó el λ óptimo mediante una validación cruzada que dividió el conjunto de datos en 5 grupos (nfold= 5). El resultado fue un $\lambda = 0.00626$, se encontró que este es el que tiene el error cuadrático medio más bajo. Sin embargo, para la regresión se utiliza λ a una desviación estándar del MMSE. Se especificó un modelo LASSO tal que las variables de

régimen de salud y jefe de hogar deban ser conservadas. Con estos parámetros, se redujo la ecuación a cinco variables dependientes:

$$\begin{aligned} \Pr[informal = 1] & \\ &= \beta_0 + \beta_1 contrato + \beta_2 meses + \beta_3 horastrab + \beta_4 edad + \beta_5 edad^2 \\ &+ \beta_6 regimen + \beta_7 jhogar + \mu_i \end{aligned}$$

En la [Figura 7](#) del Apéndice B se observa la pérdida de mínimos cuadrados en función del logaritmo de λ . Adicionalmente, se muestran las estadísticas descriptivas y la matriz de correlación de las variables empleadas.

A partir de la ecuación anterior, se plantean tres modelos: de probabilidad lineal, logit y probit. Adicionalmente, se estiman los efectos marginales y se verifica la robustez de los modelos mediante el criterio Hosmer-Lemeshow³, un pseudo-R² y la predicción sobre datos nuevos. Sin embargo, reconociendo los posibles problemas de endogeneidad debido a la misma naturaleza de la variable de interés, se plantea un modelo probit bivariado.

MPL, Probit y Logit

Los modelos MPL, probit y logit son modelos de regresión utilizados comúnmente en análisis estadísticos y econométricos. El MPL (Modelo de Probabilidad Lineal) es un modelo lineal que se utiliza para analizar la relación entre una variable dependiente binaria y una o más variables independientes continuas o discretas. El modelo estima la probabilidad de la variable dependiente basada en la relación lineal con las variables independientes.

Por su parte, los probit y logit utilizan una función logística para transformar la probabilidad lineal en una probabilidad limitada entre 0 y 1. La diferencia principal entre estos modelos radica en la función de transformación utilizada. En el modelo probit se utiliza la función de distribución normal acumulativa y en el modelo logit se utiliza la función logarítmica.

Probit bivariado

La doble causalidad entre las variables *informal* y *régimen*, significa que, bajo el marco institucional actual, no es posible inscribirse al RS siendo formal. Para ello, el modelo probit bivariado tiene en cuenta dicha endogeneidad y generaliza la estimación de tal forma que haya dos

³ Evalúa la bondad de ajuste del modelo al calcular y comparar las frecuencias esperadas contra las observadas mediante un estadístico de prueba chi-cuadrado (Hosmer & Lemeshow, 2000).

variables latentes que pueden estar correlacionadas (Cameron & Trivedi, 2005). Así, se llega al siguiente planteamiento:

$$\begin{aligned} (1) \Pr[\textit{regimen} = 1] \\ = \beta_0 + \beta_1\textit{contrato} + \beta_2\textit{meses} + \beta_3\textit{horastrab} + \beta_4\textit{edad} + \beta_5\textit{edad}^2 \\ + \beta_6\textit{ssanitario} + \beta_7\textit{jhogar} + \mu_i \end{aligned}$$

$$\begin{aligned} (2) \Pr[\textit{informal} = 1] \\ = \beta_0 + \beta_1\textit{contrato} + \beta_2\textit{meses} + \beta_3\textit{horastrab} + \beta_4\textit{edad} + \beta_5\textit{edad}^2 \\ + \beta_7\textit{jhogar} + \mu_i \end{aligned}$$

El sistema de ecuaciones propuesto tiene en cuenta que las decisiones están interrelacionadas en lugar de ser independientes, como se observaría en dos modelos probit individuales. Asumiendo que existe correlación entre las variables ($\rho \neq 0$), no existe una solución cerrada para las probabilidades (Cameron & Trivedi, 2005). Entendiendo el problema de doble causalidad entre las variables, y teniendo una correlación relevante de 0,501 ([Tabla 8](#) del Apéndice B), se procede a estimar el modelo.

Las ecuaciones indican un modelo en el que se estima la probabilidad de ser informal teniendo en cuenta el problema de endogeneidad que existe con el régimen de salud ya que en éste se incluye una variable instrumental asociada a la vivienda del individuo. La elección de la variable de servicios sanitarios fue el resultado de dos procedimientos: el primero, una selección teórica de posibles variables que pudieran explicar la participación en un régimen de salud específico al que pertenece un individuo pero que al mismo tiempo no esté asociada al fenómeno de la informalidad; el segundo, una metodología de LASSO donde se estiman dos regresiones en las que se comprueba efectivamente la relación que existe entre la variable instrumental y la variable de salud, a la vez que se verifica que no guarda relación con la informalidad.

La variable instrumental recoge información relacionada con el acceso a uno de los servicios básicos del hogar, de esta manera proporcionando información sobre cuál puede ser el régimen de salud al que pertenece un individuo pero que está completamente alejado de una explicación de por qué una persona decide ser informal.

Resultados

En la [Tabla 8](#) del Apéndice C se encuentran los resultados para los primeros tres modelos. En general, se observa que los signos son consistentes entre los modelos y no resultan contraintuitivos. La mayoría de las variables son estadísticamente significativas con un 99% de confianza. En primer lugar, como es de esperarse, el hecho de tener un contrato escrito disminuye la probabilidad de que la persona sea informal. También se puede afirmar, mediante la prueba de Wald, que el efecto conjunto de la variable jefe(a) del hogar es estadísticamente significativo. Por último, se verifica la hipótesis propuesta sobre la existencia de una relación entre la probabilidad de pertenecer al RS y la probabilidad de pertenecer al sector informal.

Se observa un R^2 del 60% para el MPL y un pseudo- R^2 de 51% tanto para el modelo probit como para el logit. Mediante los criterios de información AIC y BIC se puede afirmar que el probit es favorecido sobre el logit. Sin embargo, los resultados del criterio Hosmer-Lemeshow indican que el modelo ajusta pobremente. Un último ejercicio de predicción sobre un set de prueba con datos completamente nuevos indica una precisión del 88,04% tanto para el probit como para el logit ([Tabla 10](#) del Apéndice C).

Respecto al análisis de los efectos marginales promedio (Tabla 4), se encuentra que, contar con un contrato escrito reduce la probabilidad de ser informal en un 68 pp. Un incremento unitario en las variables de *meses*, *horastrab* y *edad*, parece no tener ningún efecto sobre la probabilidad de pertenecer al sector informal, en promedio. Sobre la relación del individuo con el jefe del hogar se observa que, los empleados del jefe del hogar son los que tienen una probabilidad mayor -en 45 pp- respecto a las demás categorías. Además, pertenecer al RS incrementa la probabilidad en 13 pp, apoyando la hipótesis sobre la relación existente en el RS y la informalidad.

Tabla 3

Efectos marginales

	Probit	Logit
<i>contrato</i>	-0,68***	-0,68***
<i>meses</i>	0,00***	0,00***
<i>horastrab</i>	-0,00***	-0,00***
<i>edad</i>	0,00***	0,00***

<i>regimen</i>	0,13***	0,13***
factor(<i>conyuge</i>)	0,03***	0,03***
factor(<i>hijo</i>)	-0,00	-0,00
factor(<i>nieto</i>)	0,01	0,01
factor(<i>otropariente</i>)	0,01*	0,01
factor(<i>empleado</i>)	0,45***	0,45***
factor(<i>pensionista</i>)	0,03	0,02
factor(<i>trabajador</i>)	0,05*	0,06**
*** p < 0,01, ** p < 0,05, *p < 0,1		

Nota. Efectos marginales promedio. Para las categorías de la variable jefe del hogar es el efecto marginal con relación al jefe de hogar. Elaboración propia.

Finalmente, se encuentra que las(os) cónyuges que pertenecen al RS mantienen una probabilidad de ser informales mayor en 13,63 pp a aquellas(os) que pertenecen al RC. Rechazando la hipótesis de que la afiliación al RC de beneficiarios del jefe del hogar incrementa la probabilidad de dichos beneficiarios de ser informales.

Respecto al modelo probit bivariado (Tabla 5), se estimó la decisión de un individuo de pertenecer o no al sector informal y de pertenecer al RS o al RC. Teniendo en cuenta que las decisiones están interrelacionadas en lugar de ser independientes, como se observaría en dos modelos probit individuales. La correlación entre ambas decisiones se refleja en ρ (*rho*), indicando una relación positiva y moderada. Dicha relación se confirma en la significancia estadística de ρ medida a través del test de Wald.

Tabla 4

Resultados modelo probit bivariado

Variables	(1) <i>informal</i>	(2) <i>regimen</i>
<i>contrato</i>	-2.466*** (0.00821)	-2.266*** (0.0122)
<i>meses</i>	0.000547*** (4.55e-05)	0.000144*** (3.73e-05)

	-	-
<i>horastrab</i>	0.000968*** (0.000201)	0.000500*** (0.000178)
<i>edad</i>	-0.00426** (0.00166)	-0.00686*** (0.00139)
<i>jhogar</i>	0.0774*** (0.0266)	0.0921*** (0.0256)
<i>conyuge</i>	0.172*** (0.0273)	-0.0629** (0.0262)
<i>hijo</i>	0.0462* (0.0268)	0.271*** (0.0260)
<i>nieto</i>	0.114*** (0.0381)	0.355*** (0.0369)
<i>otropariente</i>	0.120*** (0.0288)	0.279*** (0.0275)
<i>empleado</i>	3.181*** (0.155)	0.317*** (0.0425)
<i>pensionista</i>	0.0880 (0.128)	0.0494 (0.136)
<i>trabajador</i>	0.390*** (0.142)	0.335*** (0.119)
<i>o.P6050_9</i>	-	-
<i>edad_sq</i>	0.000176*** (1.94e-05)	0.000107*** (1.56e-05)
<i>ssanitario_1</i>		-0.668*** (0.0729)
<i>ssanitario_2</i>		0.0683 (0.0750)
<i>ssanitario_3</i>		-0.219*** (0.0816)
<i>ssanitario_4</i>		-0.0962 (0.227)
<i>ssanitario_5</i>		-0.00924 (0.119)
<i>ssanitario_6</i>		-
<i>Constant</i>	0.757*** (0.0408)	0.684*** (0.0813)
<i>rho</i>	0.363*** (0.0049)	
Observations	204,147	204,147
Standard errors in parentheses		
*** p<0.01, ** p<0.05,		
*p<0.1		
Wald test of rho=0: chi2(1) = 4612.28		
Prob > chi2 = 0.0000		

Nota. Se muestran los coeficientes estimados de las dos ecuaciones presentadas en la metodología junto con su significancia estadística, el coeficiente de correlación rho y la robustez de este mediante la prueba de Wald. Elaboración propia.

La mayoría de los coeficientes son estadísticamente significativos al 1%. Una interpretación de estos permite afirmar que, individuos jóvenes, que no cuenten con un contrato de trabajo escrito, que lleven trabajando un mayor número de meses en dicho empleo y por un mayor número de horas diarias; tienen una mayor probabilidad de pertenecer al sector informal y al régimen subsidiado.

Los efectos marginales promedio de las probabilidades conjuntas (Tabla 6) son estadísticamente significativos en su mayoría y la suma de estos a través de las cuatro probabilidades es igual a cero. Tener un contrato, reduce en 13,9 pp la probabilidad de ser informal y de pertenecer al RS. Mientras que incrementa en 34,0 pp la de ser formal y pertenecer al RC. De nuevo se observa que, los empleados de los jefes del hogar tienen una probabilidad mayor de ser informales y de pertenecer al RS en 20 pp mientras que tienen una probabilidad menor de ser formales y pertenecer al RC en 49 pp. La magnitud de estos efectos puede ser objeto de un nuevo estudio donde se profundice por las formas de empleo bajo las cuales se encuentran los empleados de los jefes de hogar, ya que se puede inferir que son relaciones en su mayoría informales donde el individuo no está protegido adecuadamente contra la enfermedad y posiblemente no realiza aportes a su pensión.

Tabla 5

Efectos marginales modelo probit bivariado

	Pr(informal=0, regimen=0)	Pr(informal=0, regimen=1)	Pr(informal=1, regimen=0)	Pr(informal= 1,regimen=1)
<i>contrato</i>	0,340***	0,139***	-0,340***	-0,139***
<i>meses</i>	-0,000***	-0,000***	0,000***	0,000***
<i>horastrab</i>	0,000***	0,000***	-0,000***	-0,000***
<i>edad</i>	0,000*	0,000*	-0,001*	-0,000*
<i>jhogar</i>	-0,009**	-0,004**	0,009**	0,004**
<i>conyuge</i>	-0,028***	-0,012***	0,028***	0,012***
<i>hijo</i>	0,001	0,000	-0,001	-0,000
<i>nieto</i>	-0,007	-0,003	0,007	0,003

<i>otopariente</i>	-0,010**	-0,004**	0,010**	0,004**
<i>empleado</i>	-0,490***	-0,200***	0,490***	0,200***
<i>pensionista</i>	-0,013	-0,005	0,013	0,005
<i>trabajador</i>	-0,052**	-0,021**	0,052**	0,021**
<i>o.P6050_9</i>	-	-	-	-
<i>edad_sq</i>	-0,000***	-0,000***	0,000***	0,000***

*** p<0.01, ** p<0.05, *p<0.1

Efectos marginales promedio. Para las categorías de la variable jefe del hogar es el efecto marginal con relación al jefe de hogar. Elaboración propia.

Conclusión

Los hallazgos mediante los modelos permiten verificar la relación estrecha que mantienen las variables de *informal* y *régimen*. Esto, dada la asociación de la protección social al estatus laboral del trabajador bajo el marco institucional actual. Las variables elegidas mediante la metodología LASSO permiten explicar en un 51% la pertenencia o no al sector informal por parte de un individuo. Adicionalmente, se encuentra que pertenecer al régimen subsidiado incrementa en 13 pp la probabilidad de pertenecer al sector informal de un individuo. Es un efecto mayor al calculado por Santa María et al. (2011) de una reducción entre 2 y 9 pp de la probabilidad de ser formal dada la pertenencia al RS. Los modelos construidos tienen una precisión del 88,04% sobre un set de datos nuevos.

Sin embargo, reconociendo el problema de endogeneidad que se presenta, dado que una persona no puede ser formal y pertenecer al RS, se propone un modelo probit bivariado que plantea dos variables dependientes que están interrelacionadas entre sí. Esto refleja de manera más clara las dos decisiones que debe tomar un individuo de pertenecer o no al sector formal y de pertenecer o no al régimen contributivo. El modelo permite afirmar que la relación existente entre las variables de interés es estadísticamente significativa, la decisión de una influye sobre la otra.

Los resultados son concluyentes y aportan a la literatura relacionada con el Sistema General de Seguridad Social en Colombia y la informalidad. Demostrando la fuerte relación existente entre estas dos variables los incentivos generados son: 1) Reducir la demanda dentro del sector formal debido a los altos CNS y al salario mínimo y, 2) Mantenerse dentro del sector informal dada la presencia de un sistema de aseguramiento gratuito. Como resultado, se tiene en un sistema de salud

insostenible y un mercado laboral segmentado, con baja productividad y protección de sus trabajadores.

Soluciones como la planteada por la Misión de Empleo 2021, de la universalidad en la protección social independientemente del nivel de ingresos de la persona, van encaminadas a solucionar la problemática planteada en el presente trabajo. Estudios posteriores podrían encaminarse en analizar la probabilidad de ser informal de un individuo dada su relación con el jefe del hogar, mediante metodologías más adecuadas.

Referencias

- Alvarado, F., Álvarez, A., Chaparro, J. C., González, C., Levy, S., Maldonado, D., Meléndez, M., Pantoja, M., Ramírez, N., & Villaveces, M. J. (2021). *Reporte ejecutivo de la Misión de Empleo de Colombia 2020 -2021*.
- Alvarado, F., Meléndez, M., & Pantoja, M. (2021). Mercados laborales fragmentados y el sistema de protección social en Colombia. In *UNDP LAC Working Paper* (Issue 14). <https://www.latinamerica.undp.org/content/rblac/es/home/library/poverty/mercados-laborales-fragmentados-y-sistemas-de-proteccion-social-.html>
- Bernal, R. (2009). The Informal Labor Market in Colombia: Identification and Characterization. *Desarrollo y Sociedad*, 145–208. http://economia.uniandes.edu.co/content/download/24070/175986/file/04_Market.pdf
- Camacho, A., Conover, E., & Hoyos, A. (2013). Effects of Colombia's Social Protection System on Workers' Choice between Formal and Informal Employment. *The World Bank Economic Review*, 1–21. <https://doi.org/10.1093/wber/lht028>
- Cameron, A. C., & Trivedi, P. K. (2005). *Microeconometrics: Methods and Applications*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511811241>
- Canavire-Bacarreza, G., Urrego, J. A., & Saavedra, F. (2017). Informalidad y movilidad en el mercado laboral: una aproximación de pseudo-panel. *Lajed*, 27, 57–75.
- Ley 100. Por la cual se crea el Sistema de Seguridad Social Integral y se dictan otras disposiciones., (1993). <https://www.funcionpublica.gov.co/eva/gestornormativo/norma.php?i=5248>
- Consejo Privado de Competitividad. (2017). *Informalidad. Una nueva visión para Colombia*. https://compite.com.co/wp-content/uploads/2017/11/CPC_Informalidad-V03.pdf
- de Soto, H. (1987). *El otro sendero: la revolución informal* (P. Instituto Libertad y Democracia (Lima (ed.)). Editorial Oveja Negra. <https://books.google.com.co/books?id=09VHAAAAYAAJ>

Departamento Administrativo Nacional de Estadística. (n.d.). *Gasto social público y privado*. Retrieved February 16, 2023, from <https://www.dane.gov.co/index.php/estadisticas-por-tema/cuentas-nacionales/cuentas-nacionales-anuales/gasto-social-publico-y-privado>

Departamento Administrativo Nacional de Estadística. (2022). *Mercado Laboral: Nueva medición de informalidad laboral*. https://www.dane.gov.co/files/investigaciones/boletines/ech/ech/Nueva_medicion_informalidad.pdf

Departamento Administrativo Nacional de Estadística. (2023). *Boletín Técnico Ocupación informal*.

Departamento Nacional de Planeación. (n.d.). *Conoce el Sisben IV*. Retrieved March 4, 2023, from https://www.sisben.gov.co/Paginas/conoce_el_sisben.aspx

Dirección de Metodología y Producción Estadística. (2019). Colombia - Gran Encuesta Integrada de Hogares - GEIH - 2019. In *Dane* (Vol. 2019). https://microdatos.dane.gov.co/index.php/catalog/599/related_materials

Gasparini, L., & Tornarolli, L. (2007). Labor Informality in Latin America and the Caribbean: Patterns and Trends from Household Survey Microdata. In *Centro de Estudios Distributivos, Laborales y Sociales*. Universidad Nacional de La Plata.

Gaviria, A., Medina, C., & Mejía, C. (2006). Evaluating the impact of health care reform in Colombia: From theory to practice. *Documento CEDE*, 1–57.

Guerrero, R., Gallego, A. I., Becerril-Montekio, V., & Vásquez, J. (2011). Sistema de salud de Colombia. *Salud Pública de México*, 53(SUPL 2), 144–155. <https://doi.org/10.1590/S0036-36342011000500003>

Hosmer, D. W., & Lemeshow, S. (2000). *Applied Logistic Regression*. John Wiley & Sons, Inc. <https://doi.org/10.1002/0471722146>

International Labour Organization (ILO). (n.d.). *Informal Employment*. Thesaurus. Retrieved February 22, 2023, from <https://metadata.ilo.org/thesaurus/936921656.html>

- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning* (Vol. 112). Springer.
- Kugler, A. D., & Kugler, M. (2003). The Labor Market Effects of Payroll Taxes in a Middle-Income Country: Evidence from Colombia. *IZA Discussion Paper*, 852. <https://doi.org/10.2139/ssrn.435380>
- Levy, S. (2008). *Good Intentions, Bad Outcomes: Social Policy, Informality, and Economic Growth in Mexico*. Brookings Institution Press. <http://www.jstor.org/stable/10.7864/j.ctt6wpfgq>
- Maloney, W. F., & Nuñez, J. (2001). Measuring the Impact of Minimum Wages Evidence from Latin America. In *Policy Research Working Paper* (Issue 2597). <http://www.nber.org/papers/w9800>
- Millán, N., & Oviedo, Y. (2009). *Encuesta social longitudinal de Fedesarrollo, etapa XIII. Cuarto seguimiento*. <http://www.repository.fedesarrollo.org.co/handle/11445/1556>
- Ministerio de Salud y Protección Social. (n.d.). *Boletín del Aseguramiento en Salud*. Retrieved March 4, 2023, from <https://www.minsalud.gov.co/proteccionsocial/Regimensubsubdiado/Paginas/coberturas-del-regimen-subsidiado.aspx>
- Ministerio de Salud y Protección Social. (2014). *Abecé movilidad entre el régimen contributivo y el régimen subsidiado*.
- Ministerio de Salud y Protección Social. (2022). *Colombia llegó al aseguramiento universal en salud al alcanzar el 99,6 %*. <https://www.minsalud.gov.co/Paginas/Colombia-llego-al-aseguramiento-universal-en-salud-al-alcanzar-el-99.6.aspx>
- Mondragón-Vélez, C., Peña, X., & Wills, D. (2010). Labor Market Rigidities and Informality in Colombia. *Serie Documentos Cede*, 11(1), 65–101. <https://www.jstor.org/stable/25800055>
- Mora, J. J., & Muro, J. (2014). Informality and minimum wages by cohort in Colombia. *Cuadernos de Economía (Colombia)*, 33(63), 469–486.

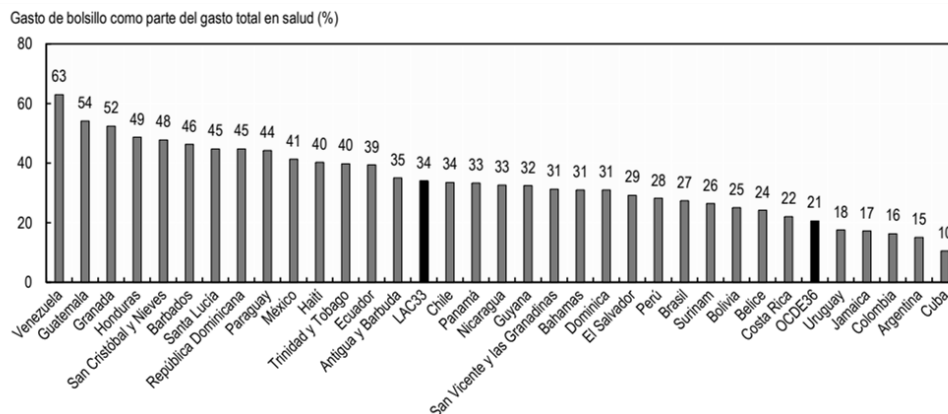
<https://doi.org/10.15446/cuad.econ.v33n63.45342>

- Organización para la Cooperación y el Desarrollo Económicos, & Banco Mundial. (2020). *Panorama de la Salud: Latinoamérica y el Caribe 2020*. rganización para la Cooperación y el Desarrollo Económicos. <https://doi.org/10.1787/740f9640-es>
- Perry, G. E., Maloney, W. F., Arias, O. S., Fajnzylber, P., Mason, A. D., & Saavedra-Chanduvi, J. (2007). *Informality: Exit and Exclusion*. World Bank Latin American and Caribbean Studies;. Washington, DC: World Bank. <http://hdl.handle.net/10986/6730>
- Ramírez Hassan, A., Cardona Jiménez, J., & Cadavid Montoya, R. (2013). The impact of subsidized health insurance on the poor in colombia: Evaluating the case of medellín. *Economía Aplicada*, 17(4), 543–556. <https://doi.org/10.1590/S1413-80502013000400001>
- Saavedra-Caballero, F., & Ospina Londoño, M. (2018). Social Assistance and Informality : Examining the Link in Colombia. *Revista de Economía Del Rosario*, 21(1), 81–120.
- Santa María, M., García, F., & Mujica, A. V. (2009). Los costos no laborales y el mercado laboral : Impacto de la reforma de salud en Colombia. In *Fedesarrollo Documentos de Trabajo* (Issue 43). <http://hdl.handle.net/11445/252>
- Santa María, M., García, F., & Mujica, A. V. (2011). El mercado laboral y la reforma a la salud en Colombia: incentivos, preferencias y algunas paradojas. In *Fedesarrollo Documentos de Trabajo*. <https://doi.org/10.33588/rn.3002.99593>
- Tibshirani, R. (1996). Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 58(1), 267–288. <http://www.jstor.org/stable/2346178>
- Vásquez, D. B., Chevreau, A., & Rijpens, J. (2022). *Transformando el empleo informal en formal: la contribución de las cooperativas en Colombia*. Informe analítico. <https://www.oecd.org/cfe/leed/social-economy/FINAL> - Informality_Colombia_Final Report.pdf

Apéndice A

Figura 3

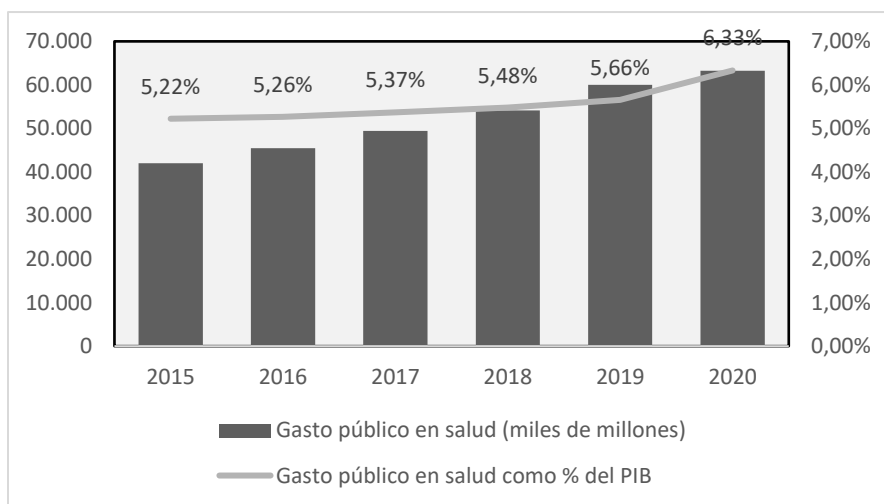
Gasto de bolsillo en salud como porcentaje del gasto total en salud en 33 países de LAC, 2017



Tomado de: *Panorama de la Salud: Latinoamérica y el Caribe 2020* (p. 12). OCDE/ Banco Mundial (2020). OCDE Publishing.

Figura 4

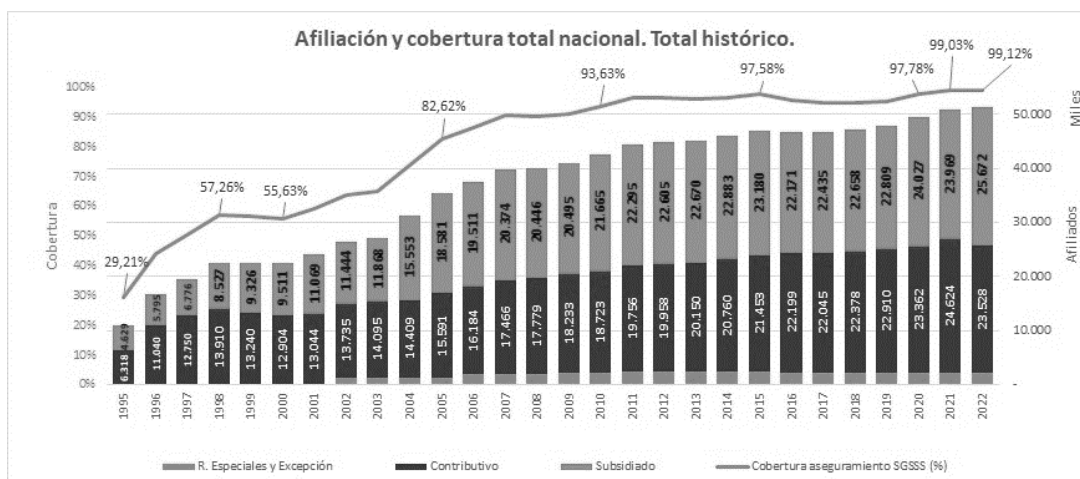
Gasto público en salud como porcentaje del PIB a precios corrientes, 2015-2020



Elaboración propia con datos de *Gasto social público y privado (SOCX)*. DANE (2021).

Figura 5

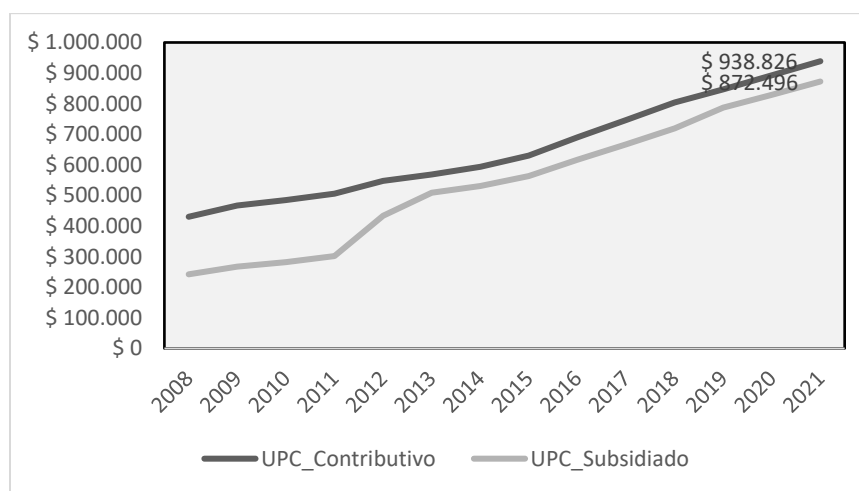
Afiliación por régimen y cobertura total sistema de salud, 1995-2022.



Tomado de: *Boletín de aseguramiento en salud*. Ministerio de Salud y Protección Social (2022).
<https://www.minsalud.gov.co/proteccion-social/Regimensubsidiado/Paginas/coberturas-del-regimen-subsidiado.aspx>

Figura 6

Unidad de pago por capitación (UPC) por régimen contributivo, 2008-2021



Elaboración propia con datos del Ministerio de Salud y Protección Social (2022).

Apéndice B

Tabla 6

Descripción de las variables empleadas

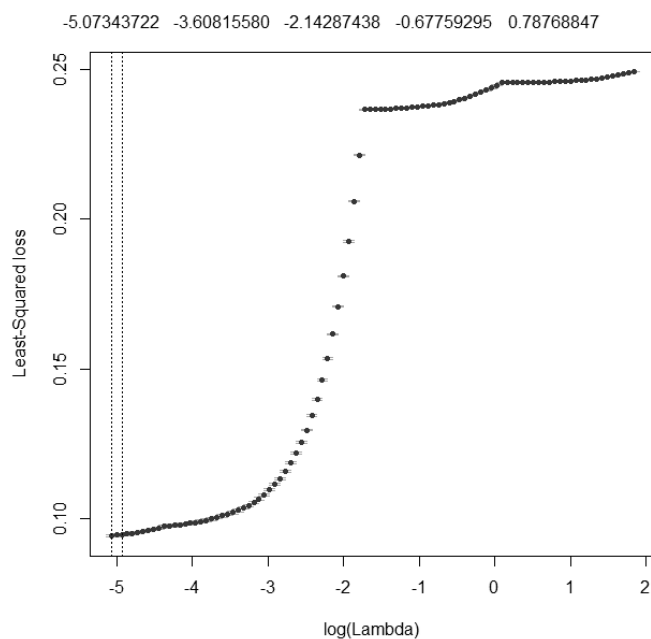
Variable	Descripción
<i>informal</i>	1 Pertenece al sector informal 0 No pertenece
<i>contrato</i>	¿Para realizar este trabajo, tiene usted un contrato escrito? 1 Sí 0 No
<i>meses</i>	¿Cuánto tiempo lleva trabajando en esta empresa, negocio, industria, oficina, firma o finca de manera continua? Meses
<i>mesestrab</i>	Número de meses trabajados en los últimos doce meses
<i>horastrab</i>	Número de horas trabajadas durante la semana pasada
<i>pension</i>	¿Está cotizando actualmente a un fondo de pensiones? 1 Sí 0 No
<i>riesgos</i>	¿Está afiliado por una empresa o individualmente a una Aseguradora de Riesgos Profesionales (ARP) (por accidentes de trabajo, enfermedad profesional etc)? 1 Sí 0 No.
<i>DPTO</i>	Departamento
<i>sexo</i>	1 Mujer 0 Hombre
<i>edad</i>	Edad
<i>edad^2</i>	Edad al cuadrado
<i>jhogar</i>	¿Cuál es el parentesco de ... con el jefe o jefa del hogar? 1 Jefe (a) del hogar 2 Pareja, esposo(a), cónyuge, compañero(a) 3 Hijo(a), hijastro(a) 4 Nieto(a) 5 Otro pariente 6 Empleado(a) del servicio doméstico y sus parientes 7 Pensionista 8 Trabajador 9 Otro no pariente
<i>ecivil</i>	Actualmente: 1 No está casado(a) y vive en pareja hace menos de dos años 2 No está casado (a) y vive en pareja hace dos años o más 3 Está casado (a) 4 Está separado (a) o divorciado (a) 5 Está viudo (a) 6 Está soltero (a)
<i>regimen</i>	¿A cuál de los siguientes regímenes de seguridad social en salud está afiliado: 1 Subsidiado? (EPS-S) 0 Contributivo (EPS)?.
<i>pagoafil</i>	¿Quién paga mensualmente por la afiliación de...? a. ... paga una parte y otra la empresa o patrón b. Le descuentan de la pensión c. ... paga la totalidad de la afiliación d. Paga completamente la empresa o patrón donde trabaja o trabajó e. No paga, es beneficiario f. No sabe, no informa
<i>educ</i>	¿Cuál es el nivel educativo más alto alcanzado por y el último año o grado aprobado en este nivel? 1. Ninguno 2. Preescolar 3. Básica primaria (1o - 5o) 4. Básica secundaria (6o - 9o) 5. Media (10o - 13o) 6. Superior o universitaria 9. No sabe, no informa
<i>ESC</i>	Años de escolaridad
<i>cuartos</i>	Incluyendo sala-comedor ¿de cuántos cuartos en total dispone este hogar?

	La vivienda ocupada por este hogar es: 1. Propia, totalmente pagada 2. Propia, la están pagando 3. En arriendo o subarriendo 4. En usufructo 5. Posesión sin título (Ocupante de hecho) ó propiedad colectiva 6. Otra
<i>vpropia</i>	
<i>nhogar</i>	Total de personas en el hogar
<i>TD</i>	Tasa de desempleo por departamento y por año
	El servicio sanitario que utiliza el hogar es: 1. Inodoro conectado a alcantarillado 2. Inodoro conectado a pozo séptico 3. Inodoro sin conexión
<i>ssanitario</i>	4. Letrina 5. Bajamar 6. No tiene servicio sanitario

Elaboración propia con datos del DANE (2019).

Figura 7

Perdida de mínimos cuadrados en función de lambda



Elaboración propia.

Tabla 7

*Matriz de correlaciones**

contrato meses horastrab edad regimen informal edad_sq

<i>contrato</i>	1	-0.104	0.104	-0.189	-0.538	-0.763	-0.208
<i>meses</i>	-0.104	1	0.042	0.502	0.052	0.127	0.503
<i>horastrab</i>	0.104	0.042	1	-0.016	-0.065	-0.086	-0.036
<i>edad</i>	-0.189	0.502	-0.016	1	0.094	0.212	0.984
<i>regimen</i>	-0.538	0.052	-0.065	0.094	1	0.501	0.110
<i>informal</i>	-0.763	0.127	-0.086	0.212	0.501	1	0.228
<i>edad_sq</i>	-0.208	0.503	-0.036	0.984	0.110	0.228	1

Elaboración propia.

*Se excluyen la variable del jefe(a) del hogar dada su construcción categórica en columnas.

Apéndice C

Tabla 8

Modelos MPL, probit y logit para la informalidad en Colombia

	MPL	Probit	Logit
(Intercept)	0,70*** (0,01)	0,49*** (0,05)	0,84*** (0,09)
<i>contrato</i>	-0,69*** (0,00)	-2,19*** (0,01)	-3,78*** (0,02)
<i>meses</i>	0,00*** (0,00)	0,00*** (0,00)	0,00*** (0,00)
<i>horastrab</i>	-0,00*** (0,00)	-0,00** (0,00)	-0,00*** (0,00)
<i>edad</i>	0,00 (0,00)	-0,00 (0,00)	-0,01** (0,00)
<i>edad^2</i>	0,00*** (0,00)	0,00*** (0,00)	0,00*** (0,00)

<i>regimen</i>	0,14*** (0,00)	0,63*** (0,01)	1,14** (0,02)
factor(<i>jhogar</i>)	0,01* (0,01)	0,05 (0,03)	0,08 (0,06)
factor(<i>conyuge</i>)	0,04*** (0,01)	0,18*** (0,03)	0,32*** (0,06)
factor(<i>hijo</i>)	0,00 (0,01)	-0,01 (0,03)	-0,04 (0,06)
factor(<i>nieto</i>)	0,01 (0,01)	0,04 (0,05)	0,05 (0,08)
factor(<i>otropariente</i>)	0,01* (0,01)	0,06 (0,03)	0,09 (0,06)
factor(<i>empleado</i>)	0,27*** (0,01)	3,19*** (0,18)	5,73*** (0,39)
factor(<i>pensionista</i>)	0,02 (0,02)	0,14 (0,16)	0,24 (0,30)
factor(<i>trabajador</i>)	0,08** (0,03)	0,30 (0,17)	0,60 (0,32)
AIC		96.504,35	96.535,54
BIC		96.652,40	96.683,59
Log Likelihood		-48.237,18	-48.252,77
Deviance		96.474,35	96.505,54
R ²	0,60	0,51 (pseudo)	0,51 (pseudo)
Adj. R ²	0,60		
Num. obs.	204.147	142.910	142.910
***p < 0,001; ** p < 0,01; *p < 0,05			

Elaboración propia.

Tabla 9

Matrices de confusión

Probit			Logit		
Referencia			Referencia		
Predicción	0	1	Predicción	0	1
0	23.322	1.525	0	23.321	1.525
1	5.801	30.589	1	5.802	30.589
Precisión	0,8803664		Precisión	0,8803501	

Elaboración propia.