

LAS FINANZAS DE LOS HOGARES MEXICANOS: ANÁLISIS CON REDES BAYESIANAS

Griselda Dávila Aragón

Escuela de Ciencias Económicas y Empresariales,
Universidad Panamericana (México)

Francisco Ortiz Arango

Escuela de Ciencias Económicas y Empresariales,
Universidad Panamericana (México)

Autor para correspondencia: fortizar@up.edu.mx

Agustín I. Cabrera Llanos

Unidad Profesional Interdisciplinaria de Biotecnología,
Instituto Politécnico Nacional (IPN, México)

Recibido el 03 de octubre de 2020; aceptado el 14 de abril de 2021.

RESUMEN

El bienestar de los hogares está ligado en gran parte al desarrollo de los mercados financieros. El estudio de las finanzas de los hogares analiza las formas en que estos utilizan instrumentos financieros para satisfacer sus necesidades y objetivos; este análisis representa un gran desafío debido a la escasa información estadística y la interrelación entre las variables consideradas. En este trabajo, pionero en el uso de las redes bayesianas en este campo, utilizamos de manera conjunta las finanzas tradicionales y las conductuales. Medimos la probabilidad de prevalencia de estabilidad financiera de los hogares en México; obtenemos un resultado base y posteriormente, al generar distintos escenarios, descubrimos que las variables más determinantes son el manejo del crédito y la conformación de los hogares. Estos resultados subrayan la importancia de promover iniciativas de educación financiera en los distintos niveles, modalidades y subsistemas educativos.

Palabras clave: finanzas del hogar, mercados financieros, redes bayesianas.

Clasificación JEL: C11, C67, D14.

<http://dx.doi.org/10.22201/fe.01851667p.2021.317.77127>

ABSTRACT

In today's economy, the well-being of households is considered to be linked mainly to the development of financial markets. The field of household finance analyzes how households use financial instruments to satisfy their needs and achieve objectives. This analysis represents a significant challenge due to scarce statistical information and the interrelation among the variables involved. We follow two aspects: Traditional and behavioral finance. This paper pioneers the use of Bayesian networks in the field. A model measuring the probability of prevalence of financial stability of households in Mexico is used; a baseline result is obtained and then, while generating different scenarios, we discover that credit management and household composition are the most determining variables. These results underscore the importance of promoting different financial education initiatives at different educational levels, modalities, and subsystems.

Keywords: Household finance, financial markets, Bayesian networks.

JEL Classification: C11, C67, D14.

1. INTRODUCCIÓN

De acuerdo con el Instituto Nacional de Estadística y Geografía de México (INEGI), un hogar está formado por una o más personas que residen habitualmente en la misma vivienda que pueden ser parientes o no y sostenerse con un ingreso común, utilizado principalmente para alimentos. Según la misma definición, el jefe del hogar será la persona reconocida así por todos sus integrantes, sin particularidad de género (INEGI, 2018).

Las finanzas del hogar (FH) se enfocan en el estudio normativo y positivo de cómo sus miembros utilizan el mercado financiero para alcanzar sus objetivos (Gogolin, Dowling y Cummins, 2017); también podemos decir que las FH se dedican a estudiar las formas en las que los jefes del hogar utilizan instrumentos financieros que les permitan alcanzar sus objetivos (Badarinza, Campbell y Ramadorai, 2016). La planeación financiera en los hogares se realiza sobre horizontes de largo

plazo, pero finitos y con limitantes en su capacidad de solicitar préstamos (Brown, Gaino y Taylor, 2013). Para llevar a cabo el estudio de las FH, enfocado en el caso de México, se pueden seguir dos vertientes: la primera basada en la teoría de las finanzas tradicionales, que parte de decisiones racionales (comportamiento ideal) y supone que los tomadores de decisiones tienen cierto grado de educación financiera y procuran maximizar su beneficio. La segunda vertiente se basa en las finanzas conductuales, la cual considera que las personas no toman las decisiones de manera óptima necesariamente e implica contar con datos reales, que no siempre están disponibles (Georgarakos, Haliassos y Pasini, 2014). La brecha entre ambas vertientes suele ser amplia y difícil de conciliar.

En esta investigación pretendemos conciliar ambas vertientes aplicando la metodología de redes bayesianas (RB); así que partimos de considerar que los jefes del hogar no siempre toman decisiones de manera racional y óptima, dado el supuesto que tienen acceso a información financiera suficiente. El estudio se lleva a cabo en hogares mexicanos. Un dato relevante del presente trabajo es que es pionero en el uso de un modelo basado en redes bayesianas para analizar este fenómeno. El antecedente con mayor similitud es D'Urso y Vitale (2020), que se propone el uso de redes bayesianas para medir los indicadores básicos y compuestos del bienestar equitativo y sostenible y que afirma que esto se puede hacer gracias a la capacidad de las RB para modelar estructuras de dependencia multivariante complejas, además de que permiten describir las relaciones entre indicadores pertenecientes a diferentes dominios. Aprovechamos estas mismas propiedades de las RB para conjuntar las dos vertientes de estudio de las FH mencionadas previamente.

El presente trabajo está dividido en cinco secciones: en la siguiente describimos de manera detallada la información estadística relativa a aspectos demográficos, sociales, culturales y económicos de la población y de los hogares en México. En la sección tres presentamos una breve descripción de la metodología de las redes bayesianas, que utilizamos como herramienta de análisis; también construimos y aplicamos el modelo para obtener los resultados en forma de un diagrama que muestra los valores del estado base de la salud financiera de los hogares en México. En la cuarta realizamos cambios extremos a los valores asignados a algunas de las variables más sensibles con objeto de generar tres distintos escenarios con probabilidades de ocurrencia considerable, en los que

se distingue que el manejo adecuado del crédito y la conformación del hogar son las dos variables que afectan de manera más significativa a las finanzas en los hogares de México. Estos hallazgos pueden orientar el diseño y la aplicación de políticas públicas referentes a la educación financiera de la población y la planeación de programas gubernamentales destinados a grupos económicamente vulnerables. La última sección son las conclusiones.

2. CONTEXTO DEMOGRÁFICO

Según el INEGI, en México había 125 millones de personas en 2018, el 51.4% eran mujeres y 48.6% hombres. El 75.5% de la población habitaba en localidades urbanas¹ y el resto en zonas rurales;² con un promedio de 3.6 integrantes en 34.7 millones de hogares, en los que 2.4 integrantes percibían ingresos, pero sólo 1.7 tenían un empleo formal. Asimismo, el 71.3% de los hogares los encabezaba un hombre y el resto una mujer, con edad promedio de 49.8 años. En el 76% de los casos, el jefe de familia se encontraba en un rango de edad entre 18 y 50 años (INEGI, 2018).

El *Censo de Población y Vivienda 2020* (INEGI, 2020) indica que los hogares familiares componen el 87% de los hogares, el 71.3% son nucleares, es decir, están formados únicamente por parejas o uno solo de los padres, con o sin hijos; el 27.9% son hogares ampliados, que son hogares nucleares más otros parientes de una de las parejas y, finalmente, el 0.8% son hogares compuestos, esto es, formados por parejas o un solo padre, con o sin hijos o familiares y adicionalmente personas que no tienen ningún parentesco con el jefe del hogar.

La *Encuesta Nacional de Ingresos y Gastos de los Hogares* (ENIGH) [INEGI, 2018] muestra que en 2018 el ingreso promedio total trimestral por hogar fue de 53 012 pesos, constituido por un ingreso corriente de 49 610 pesos más 3 402 pesos correspondientes a percepciones financieras y de capital; de los 49 610 pesos, el 67.3% corresponde a la remuneración por el trabajo desempeñado por los integrantes del hogar, 15.4% por transferencias, 11.4% por alquiler de la vivienda, 5.9% por renta de la

¹ Localidades con más de 2 500 habitantes

² Localidades con menos de 2 500 habitantes

propiedad y 0.08% por otros ingresos corrientes.³ Si agrupamos a todos los hogares (urbanos y rurales) por deciles, el rango de ingresos diarios por hogar oscila entre 101 pesos para el primer decil y 1 853 pesos para el décimo decil.

Por otro lado, en las localidades urbanas el ingreso corriente promedio diario por receptor del hogar en el primer decil fue de 46 pesos y en el décimo decil de 804 pesos. En el caso de las localidades rurales dichos ingresos fueron de 37 pesos y 709 pesos respectivamente. En términos de género, existe una gran diferencia entre las percepciones de hombres y mujeres. En la población no indígena, el género masculino tiene un ingreso promedio trimestral de 19 590 a 30 107 pesos y los ingresos de las mujeres varían de 7 774 a 16 195 pesos. Para la población considerada indígena, los ingresos son aún menores. En términos de nivel de escolaridad, el ingreso promedio trimestral para personas con primaria terminada es de 8 527 pesos y para aquellas con estudios de posgrado es de 86 880 pesos.

El gasto promedio total trimestral por hogar fue de 47 205 pesos; el gasto corriente monetario⁴ promedio trimestral por hogar fue de 31 913 pesos y el no monetario⁵ de 8 673 pesos; el resto corresponde a erogaciones financieras y de capital.⁶ Un hogar destina en promedio el 35% del ingreso a alimentos, bebidas y tabaco; el 20% a transporte y comunicaciones y el resto a servicios de educación y esparcimiento. Cabe destacar que, sin distinción de localidad, el menor gasto es el dedicado a la salud. En las zonas rurales el gasto corriente monetario promedio trimestral fue de 19 938 pesos y en las zonas urbanas fue 1.78 veces mayor.

La *Encuesta Nacional de Ocupación y Empleo* (ENOE) [INEGI, 2019] indica que, durante el cuarto trimestre de 2018, el 54.34% de la pobla-

³ La suma es mayor al 100% debido al redondeo en las cifras.

⁴ Ingreso destinado durante el periodo de referencia para la adquisición de productos y servicios de consumo final y privado, el gasto efectuado en dinero o por la compra de productos y/o servicios que fueron pagados, donados y/o regalados como una transferencia corriente a personas o instituciones ajenas al hogar.

⁵ Estimación con base en el valor de mercado al menudeo de los productos y servicios de consumo final y privado. Se clasifican en autoconsumo, pago en especie, regalos y estimación del alquiler de la vivienda.

⁶ Aquellas transacciones destinadas a la compra de bienes muebles e inmuebles, objetos valiosos, activos físicos que modifican el patrimonio del hogar.

ción económicamente activa en México laboraba en el sector informal y la tasa de desocupación era de 3.4%. Por otro lado, el 42.42% de los hogares pertenecía a la clase media, lo que implicaba tener computadora e internet, el gasto en alimentos era menor al 35% del ingreso y al menos uno de sus integrantes era asalariado formal del sector privado. El jefe de familia tenía educación media superior o más, la vivienda era propia o financiada con recursos de la familia o algún crédito. El 55.08% de los hogares era de clase baja y solamente el 2.5% pertenecía a la clase alta.

Debido a que la principal fuente de ingresos de los hogares mexicanos es la actividad laboral de sus miembros, perder el empleo desestabiliza las FH. A consecuencia de la pandemia por COVID-19 se ha perdido una gran cantidad de empleos. El Centro de Estudios Espinoza Yglesias (CEEY, 2020) afirma que, hasta la primera quincena de abril de 2020, entre 5.2 y 8.1 millones de personas perdieron su empleo en México debido a las medidas sanitarias impuestas que implicaron el cierre de negocios y empresas. Esto propició que en 37.7% de los hogares al menos un integrante perdiera su fuente de ingresos y estima adicionalmente que en 30.1% de los hogares otro miembro podría perder el empleo.

El Consejo Nacional de Evaluación de la Política de Desarrollo Social (CONEVAL) ubica a una persona en situación de pobreza cuando tiene al menos una carencia social en alguno de estos seis indicadores: rezago educativo, acceso a servicios de salud, acceso a la seguridad social, calidad y espacios de la vivienda, servicios básicos en la vivienda y acceso a la alimentación y su ingreso es insuficiente para adquirir los bienes y servicios que se requieren para satisfacer sus necesidades alimentarias y no alimentarias. Además, distingue entre pobreza moderada y extrema; esta última se presenta cuando una persona tiene tres o más carencias de las seis posibles y adicionalmente se encuentra por debajo de la línea de bienestar mínimo. Las personas en pobreza extrema disponen de un ingreso tan bajo, que aun si lo dedicasen por completo a la adquisición de alimentos no podrían adquirir los nutrientes necesarios para tener una vida sana (CONEVAL, 2019).

El CONEVAL sostiene que en 2018 había 52.4 millones de personas en situación de pobreza, y de ellas 21 millones en pobreza extrema (CONEVAL, 2019). Ante la inminente pérdida de empleos debida a la pandemia de COVID-19, medidas que favorezcan las FH y permitan mejorar las condiciones de vida para la población adquiere relevancia.

En el presente trabajo empleamos la metodología de redes bayesianas para diagnosticar el estado de una adecuada administración de los ingresos y gastos del hogar en los hogares mexicanos, utilizamos la definición dada por (Weida *et al.*, 2020): “La salud financiera, debe ser entendida como la capacidad de la persona para administrar sus gastos, prepararse y recuperarse de las crisis financieras, tener una deuda mínima y la capacidad de generar riqueza, es la base de todas las facetas de la vida diaria, como asegurar los alimentos y pagar la vivienda”. Una ventaja de utilizar las RB es que permiten emplear datos estadísticos y, al mismo tiempo, trabajar con la incertidumbre de estimaciones subjetivas elaboradas a partir de información proveniente de expertos ante la ausencia de datos estadísticos; de esta manera, las estimaciones se van ajustando conforme se registran más datos observados.

Un reto para analizar las FH es la falta de información sobre el comportamiento de las inversiones de las familias en México. Medir el comportamiento financiero de los hogares es complejo y los modelos financieros tradicionales no reflejan la realidad que enfrentan los hogares (Campbell, 2006). En general los hogares tienen más de una fuente de ingresos, inversiones o ahorros en diferentes instituciones bancarias, préstamos informales, gastos hormiga y otros factores, que aunados a la confidencialidad de la información imposibilitan tener datos estadísticos confiables para esa variable. A partir de 2018 la Comisión Nacional Bancaria y de Valores (CNBV) elabora la *Encuesta Nacional de Inclusión Financiera* (ENIF), para medir la inclusión financiera⁷ y publica sus resultados mediante el *Reporte Nacional de Inclusión Financiera* (RNIF). La restricción observada en tal encuesta es que la información se presenta por individuo y no por hogar.

De acuerdo con la CNBV (2019), en 2018 el 68% de las personas entre 18 y 70 años tenía al menos un producto financiero,⁸ 72% del género

⁷ Acceso y uso de servicios financieros formales con una regulación apropiada que garantice esquemas de protección al consumidor y promueva la educación financiera para mejorar las capacidades financieras de todos los segmentos de la población (CNBV, <<https://www.gob.mx/cnbv/acciones-y-programas/inclusion-financiera-25319>>, consultado el 08/06/2020).

⁸ Con un banco o institución financiera (nómina o ahorro), crédito formal (tarjeta de crédito, crédito de vivienda personal u otros), seguro o cuenta de Ahorro para el Retiro (Afore).

masculino. El 24% poseía un producto financiero, el 27% dos o tres y el 17% cuatro o más de ellos. El 47% de las personas con un producto financiero tenía una cuenta bancaria, el 40% una cuenta de ahorro para el retiro o pago de nómina. El 31% tenía un crédito, de estos el 61% era una tarjeta de crédito departamental o de tienda de autoservicio, el 34% tarjeta de crédito bancaria y el 5.4% un crédito de vivienda.

Para efectos de créditos, la edad de los integrantes del hogar es fundamental, las parejas jóvenes tienen un horizonte de tiempo más largo; para parejas maduras su horizonte es mucho menor, por lo que son más propensas a tener ahorros para la vejez y ser más reservados con sus inversiones. (Bodie, Merton y Samuelson, 1992; Campbell y Viceira, 2002). Según Polkovnichenko (2005), la mayoría de los hogares no posee activos financieros, un escaso porcentaje tiene cuentas de jubilación y en el menor de los casos poseen inversiones en el mercado de capitales. Debido a que un alto porcentaje de hogares en México obtiene sus ingresos de la economía informal, que están libres de gravamen, no consideramos la variable impuestos en la presente investigación.

3. LAS REDES BAYESIANAS

Las RB son una estructura matemática que se utiliza para modelar las relaciones de dependencia entre variables aleatorias, éstas permiten propagar sistemáticamente el impacto de una acción sobre las variables consideradas de acuerdo con sus probabilidades de interacción y relaciones de dependencia, dando origen a resultados aleatorios. En la representación gráfica de estos modelos hay dos elementos fundamentales, los nodos y las aristas. Los nodos representan las variables de interés y las aristas son los enlaces causales o de influencia entre las variables (Madsen y Kjaerulff, 2008).

En una RB las variables pueden ser cualitativas o cuantitativas, conectadas entre ellas mediante aristas, tienen además dos tipos de nodos, el “nodo padre” que es el nodo antecesor y el sucesor, denominado “nodo hijo”. Los nodos padre tienen asignada una probabilidad marginal, toda vez que no dependen de ninguna otra variable, y los nodos hijo tienen probabilidades condicionales, al ser dependientes de los nodos padre (Chan *et al.*, 2018). Cada nodo hijo tiene asociada una distribución de probabilidad condicional que sólo depende de sus padres, la cual se des-

cribe mediante una tabla de probabilidad, una distribución estadística o una función parametrizada.

También podemos conceptualizar a una RB como la representación de una declaración simple con independencia condicional. Es decir, que cada variable que la integra es independiente de sus no descendientes en el gráfico, dado el estado de sus padres. Esta propiedad es aprovechada para reducir de manera significativa el número de parámetros requeridos en la caracterización de la distribución de probabilidad conjunta de las variables consideradas. Esta reducción permite realizar un cálculo eficiente de las probabilidades posteriores dados los valores iniciales (Spirtes, Glymour y Schienes, 1993).

Las RB pertenecen a una clase de modelos gráficos probabilísticos conocida como gráfico acíclico dirigido (GAD). La estructura de un GAD garantiza que no existan nodos que puedan ser su propio padre o su propio hijo, esta condición es de vital importancia para el cálculo de la probabilidad conjunta (CPC) de una colección de nodos mediante la factorización de la probabilidad de cada nodo (Pearl y Russel, 2001).

Además de la estructura GAD, que a menudo se considera como la parte “cualitativa”, es necesario especificar los parámetros “cuantitativos” del modelo. Los parámetros se describen de manera que la distribución de probabilidad condicional en cada nodo dependa sólo de sus padres. Para variables aleatorias discretas, esta probabilidad condicional normalmente se representa mediante una tabla, enumerando la probabilidad local de que un nodo hijo tome cada uno de los valores factibles para cada combinación de valores de sus padres. La dependencia se puede determinar a partir de la distribución de una colección de variables mediante las tablas de probabilidades condicionales locales.

La tabla de probabilidad de un nodo es un arreglo que lista un conjunto de valores de probabilidad condicional que modelan la relación de incertidumbre entre este nodo y los nodos con los que interactúa, ya sean nodos padre o nodos hijo, junto con cualquier incertidumbre presente en esa relación. Por ejemplo, sean Y y X dos variables aleatorias, supongamos que una arista está dirigida desde Y , el nodo padre, a X , el nodo hijo, la arista representa la influencia de la variable Y sobre la variable X , la cual se cuantifica mediante la probabilidad condicional asociada (Pinheiro *et al.*, 2003). Las RB tienen la virtud de ser matemáticamente rigurosas e intuitivamente comprensibles, y permiten establecer una

representación y medio de cálculo efectivos de la CPC sobre un conjunto de variables aleatorias (Pearl, 1988).

Los métodos bayesianos constituyen una de las técnicas más utilizadas en los problemas de inteligencia artificial, el aprendizaje automático y la minería de datos (Mitchell, 1997). Lo anterior en virtud de que constituyen un método válido y práctico para realizar inferencias con los datos de que se disponga, lo que implica inducir modelos probabilísticos que, una vez calculados, se pueden utilizar con otras técnicas de minería de datos (Beltrán, Muñoz y Muñoz, 2014).

Lo anterior sitúa a las RB como una herramienta poderosa que permite incluir simultáneamente la opinión subjetiva de expertos y datos objetivos observados en modelos de causa-efecto. Al combinar los datos estadísticos con los datos cualitativos proporcionados por expertos, es posible construir una estructura causal subyacente al tema de análisis. Las RB se han erigido como una herramienta capaz de integrar criterios y aspectos cualitativos subjetivos con datos cuantitativos, facilitando la comprensión causal en las estimaciones (Guidici, 2004). Esta propiedad nos permitió conciliar en este trabajo los enfoques de las finanzas tradicionales y las finanzas conductuales para analizar las FH.

Su mayor ventaja es que facilitan razonar en condiciones de incertidumbre, al combinar las bondades de la representación visual intuitiva con una base matemática de probabilidad bayesiana (Uusitalo, 2007), lo que permite articular dependencias entre diferentes variables y propagar sistemáticamente el impacto de la evidencia sobre las probabilidades de resultados aleatorios (Neil, Fenton y Tailor, 2005). La clave para el diseño exitoso de una RB es la descomposición adecuada del dominio del problema en un conjunto de proposiciones causales o condicionales. Por esta razón procedemos a dividir el problema y tomamos las especificaciones parciales del modelo que sean significativas, con base en el dominio y conocimiento del experto (Supatgiat, Kenyon y Heusler, 2006).

Una vez determinadas las variables a considerar, ya sean cualitativas o cuantitativas, asignamos a cada variable una tabla de probabilidad, la cual se construye usando datos históricos y, en ausencia de estos, solicitando a un experto que nos proporcione una serie de estimaciones, mismas que estarán basadas en su conocimiento y experiencia (Cowell *et al.*, 1999). Una RB es un GAD que representa una distribución de probabilidad conjunta sobre un conjunto de variables aleatorias V , donde la red puede

definirse mediante el par ordenado $B = (G, Q)$, G es el GAD, cuyos nodos X_1, X_2, \dots, X_n representan variables aleatorias y cuyas aristas representan la dependencia directa entre estas variables. El gráfico G codifica los supuestos de independencia, por lo que cada variable X_i es independiente de sus variables no descendientes dados sus padres en G . La componente Q denota el conjunto de parámetros de la red. Este conjunto contiene el parámetro $\theta_{x_i|\pi_i} = P_B(x_i|\pi_i)$ para cada realización $x_i \in X_i$ condicionada a π , el conjunto de padres de X_i en G . En consecuencia, B define una CPC única sobre V (Ben-Gal, 2007). Con lo cual, el valor de la probabilidad condicional se calcula mediante la expresión:

$$P_B(X_1, X_2, \dots, X_n) = \prod_{i=1}^n P_B(X_i|\pi_i) = \prod_{i=1}^n \Theta_{X_i|\pi_i} \quad [1]$$

Expresión que utilizaremos para calcular las probabilidades en cada nodo de la RB, como se verá más adelante.

3.1. Metodología de redes bayesianas

Iniciaremos con la construcción del modelo, para lo cual procedemos de acuerdo con el orden siguiente:

1. Definimos el dominio del problema en el que especificamos el propósito de la red, es decir, identificamos la causalidad entre las variables que impactan a las FH. Una vez definido el propósito de la red, identificamos las variables o nodos importantes para delimitar el dominio del problema. El proceso de identificación de las variables para definir los nodos es la parte medular del diseño de la red, esto requiere un conocimiento profundo del proceso a modelar, y así poder descomponerlo en todas sus fases, seleccionar los nodos y establecer las relaciones entre ellos.
2. Para identificar las variables más importantes consultamos a los expertos en el proceso. Identificamos y definimos 10 variables y sus posibles valores (véase el cuadro 1). Los expertos consultados son colaboradores de empresas y reguladores financieros que tienen la experiencia e información sobre la operación y administración de proyectos en FH. Por confidencialidad, su centro de trabajo y nombres no pueden ser revelados. A través de las entrevistas realizadas a estos expertos nos fue

posible asignar valores a las probabilidades *a priori* para los diferentes nodos. La confianza en la precisión de las sugerencias de los expertos es indispensable para utilizar este método. A su vez, los expertos basan sus opiniones en los datos de que dispongan, pero matizados con base en su experiencia y conocimiento para proyectarlas hacia eventos futuros.

Cuadro 1. Definición de las variables (nodos) y sus dimensiones (valores) posibles

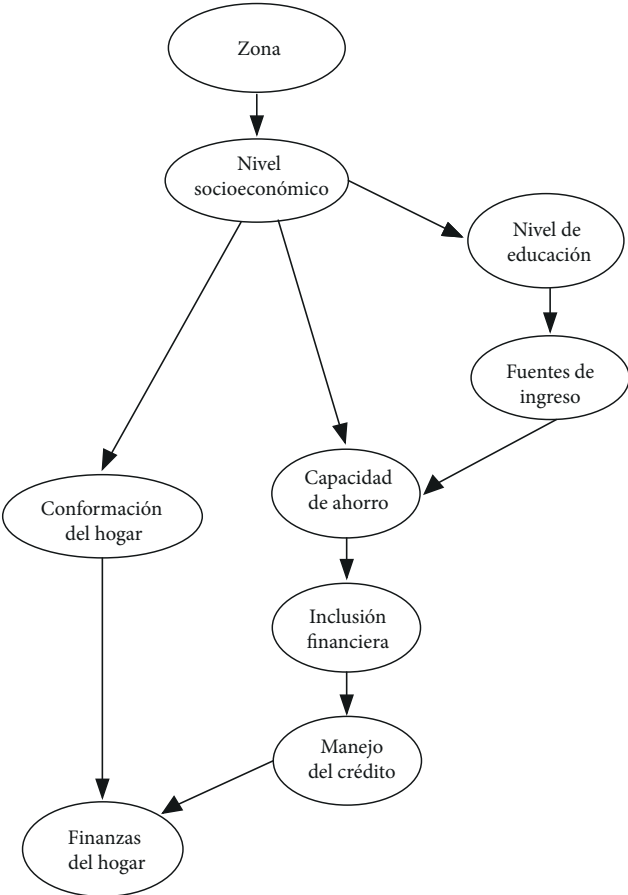
	Variable	Dimensiones
1.	Nivel socioeconómico	Alta Media Baja
2.	Conformación del hogar	Nuclear/Monoparental Ampliado Compuesto
3.	Nivel de educación	Profesionales Media superior Básica ¹
4.	Inclusión financiera	Un producto Dos o tres productos Cuatro o más productos
5.	Zona	Rural Urbana
6.	Fuente de ingreso	Formal Informal
7.	Capacidad de ahorro	Alta Media Baja
8.	Manejo del crédito	Adecuado No adecuado
9.	Finanzas del hogar	Estables No estables

Nota: 1/ la educación básica comprende desde preescolar hasta secundaria.

Fuente. elaboración propia.

Con las variables seleccionadas y la información proporcionada por los expertos consultados, elaboramos el modelo gráfico de la RB, definimos las relaciones de causalidad entre los nodos y así obtuvimos una red que consta de un nodo padre: Zona y siete nodos hijo: Nivel socioeconómico, Nivel de educación, Fuentes de ingreso, Capacidad de ahorro, Conformación del hogar, Inclusión financiera y Manejo del crédito. Con este modelo procedemos a evaluar la variable FH, que mide la probabilidad de contar con finanzas estables en un hogar. Las relaciones resultantes entre los nodos se representan en el diagrama 1.

Diagrama 1. Modelo gráfico de la red bayesiana de las finanzas del hogar



Fuente. elaboración propia empleando el software GeNIe.

Una vez definido el modelo gráfico de la RB, lo validamos con el grupo de expertos en FH mediante un *focus group*, lo que nos permitió validar las probabilidades *a priori* utilizadas como primer insumo.

3.2. Cuantificación del estado base de la red bayesiana

Para cuantificar la RB por parte de los expertos, consideramos que cada nodo en la red tiene asociada una tabla de probabilidad condicional que determina el nivel de interrelación de los nodos. Estas probabilidades las propusimos a partir de la información obtenida del INEGI y de la CNBV y fue validada por el grupo de expertos en FH con base en el grado de creencia⁹ respecto al comportamiento del parámetro aleatorio.

Los datos estadísticos obtenidos permiten iniciar el proceso de cuantificación; sin embargo, no existía información de las probabilidades condicionales para cada nodo, por lo que aprovechamos una propiedad de la estadística bayesiana que permite tratar a los parámetros como variables aleatorias. De este modo, las afirmaciones en relación con las características de una variable no dependen sólo de datos estadísticos, sino también de cualquier conocimiento *a priori* que se tenga del estadístico, en este caso la opinión subjetiva de los expertos.

Debido a lo laborioso de los cálculos necesarios para determinar los valores de cada fase de la propagación entre las interacciones de los nodos, sus ponderaciones y las probabilidades asociadas a cada proceso, en la presente investigación utilizamos una versión libre del software GeNIe para calcularlos. A partir de la cuantificación de la red es posible calcular la probabilidad de contar con finanzas estables en un hogar. Para realizar estos cálculos iniciamos con probabilidades *a priori*, al ejecutar el algoritmo de la RB obtenemos las probabilidades *a posteriori* para cada nodo de la red, las cuales son mostradas en el diagrama 3, cuyos valores proporcionan la probabilidad de la existencia de finanzas estables en el hogar.

⁹ Los grados de creencia representan formalmente la fuerza (intensidad) con la que aceptamos la verdad de varias proposiciones. Cuanto mayor sea el grado de creencia de un agente para una proposición particular, mayor será su confianza en la verdad de esa proposición (Huber, 2009).

Para ilustrar la forma en que GeNIe calcula los valores de cada nodo calculamos, a manera de ejemplo, el valor en uno de ellos, para lo cual consideramos los siguientes valores: la proporción de la población urbana (U) corresponde al 75.5% y el 24.5% a la población rural (R). Siendo A_1 el nivel socioeconómico de la familia, la probabilidad de que el nivel socioeconómico de la familia sea alto, dado que pertenece a la población urbana, es de 3%, y que sea alto, dado que pertenece a la población rural, es de 0.614%. Ambos valores se obtienen promediando los valores asignados por los expertos. A partir de esto efectuamos el cálculo del valor de la probabilidad de cada elemento del nodo hijo:

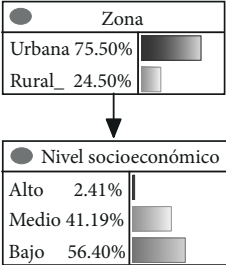
$$\begin{aligned}
 P(U) &= 75.5\% \\
 P(R) &= 24.5\% \\
 P(A_1|U) &= 0.03 \text{ y } P(A_1|R) = 0.00614
 \end{aligned}$$

De este modo, la probabilidad de pertenecer a la clase socioeconómica alta es:

$$\begin{aligned}
 P(A_1) &= P(U)P(A_1|U) + P(R) P(A_1|R) \\
 &= (0.755)(0.03) + (0.245)(0.00614) = 0.0241543
 \end{aligned}$$

De manera análoga se calculan las probabilidades de los otros niveles socioeconómicos con la información proporcionada por los expertos. Para efectos de la presentación de los resultados, el valor de 0.0241543 se redondea a 0.0242, como aparece en el diagrama 2 generado por el software GeNIe.

Diagrama 2. Valores del primer nodo hijo



Fuente. elaboración propia empleando el software GeNIe.

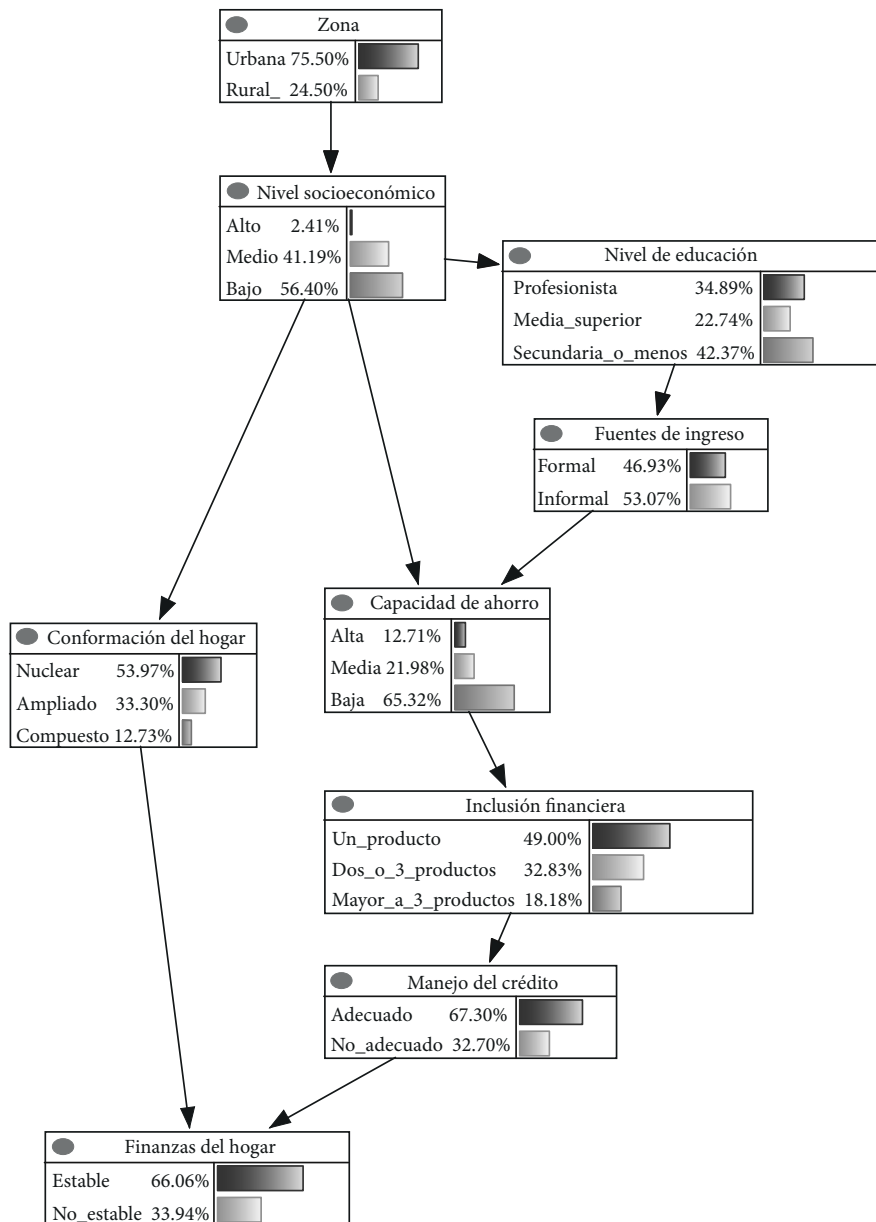
Para los demás nodos, los cálculos se llevaron a cabo de manera análoga, utilizando GeNIe, como se indicó antes, debido a lo laborioso de éstos al aumentar las relaciones de dependencia entre las variables. En los resultados mostrados puede haber discrepancias en las cifras por efectos del redondeo (véase el diagrama 3).

Al analizar los valores de la red en el escenario inicial o base (redondeando los porcentajes) tenemos los siguientes resultados:

- El 75% de los hogares en México se localiza en zonas urbanas, con probabilidad de 41% de que su nivel socioeconómico sea medio y 56% de que pertenezca al nivel socioeconómico bajo.
- Con probabilidad del 35% que el jefe de familia posea estudios profesionales y con 42% de que su nivel de estudios sea básico o menor.
- Los ingresos del hogar se obtienen de una fuente de ingresos formal con probabilidad del 47% y con el 53% de que su origen sea de fuentes informales.
- Respecto a la capacidad de ahorro, con probabilidad de 13% el ingreso del hogar permite un ahorro alto, el 22% ahorro medio y en el 65% el ahorro es bajo o nulo.
- En lo referente al acceso a productos financieros, con probabilidad de 49% los hogares tienen acceso a un producto solamente, con 33% tienen acceso a dos o tres productos y con 18% a cuatro productos o más.
- El acceso a productos financieros incrementa el acceso a crédito por parte de las instituciones financieras; con probabilidad del 67% el manejo del crédito en estos hogares será adecuado y el 33% no lo será.
- Respecto a la conformación del hogar, el 54% de los hogares lo integran ambos padres o uno e hijos (nuclear), el 33% son ampliados y el 13% compuestos.
- En el escenario antes descrito, la probabilidad de un balance equilibrado con una buena administración de los ingresos y gastos del hogar o finanzas estables es de 66% y de finanzas no estables el 34%.

Los resultados anteriores son relevantes para elaborar un perfil de clientes potenciales del sistema bancario comercial, así como para la planeación de programas gubernamentales destinados a apoyar grupos económicamente vulnerables.

Diagrama 3. Modelo gráfico de probabilidades *a posteriori* del estado inicial o base de las FH de los hogares en México



Fuente. elaboración propia empleando el software GeNIe.

4. ANÁLISIS DE ESCENARIOS

A continuación llevaremos a cabo un análisis de diferentes escenarios que podrían presentarse con mayor probabilidad. Estos describen circunstancias que pueden ocurrir en un futuro y permiten explorar diferentes opciones (Banuls y Salmeron, 2007). Una de las mayores ventajas de emplear RB es que posibilitan desarrollar de forma rápida los análisis de diversos escenarios, considerando la interacción causal entre las variables que conforman la red. Los datos ingresados en cada nodo se propagan por toda la red en dirección descendente y de esta manera se actualiza la distribución de probabilidad marginal de cada nodo y obtenemos diferentes probabilidades de prevalencia de salud financiera en los hogares mexicanos.

Para realizar los análisis de escenarios considerados, empezamos por evaluar en la RB a una de las variables con el valor de probabilidad más alto posible; esto se interpreta como un choque exógeno que elimina la relación directa con el nodo inmediato superior y provoca que la arista que los une en el modelo desaparezca. Llevar a cabo estas asignaciones de valores a distintas variables requiere realizar previamente un análisis de sensibilidad que permita identificar las variables que incrementen sustancialmente la probabilidad de tener finanzas estables en el hogar y aplicar los recursos y esfuerzos de mejora en ellas, con la restricción de que en algunos casos esto implicaría un cambio estructural en la política del país. Los resultados obtenidos con los nuevos supuestos se comparan con los del estado base y esto nos proporciona un panorama de las condiciones de cada estado. A continuación presentamos el desarrollo de este análisis.

La zona en que se asientan los hogares, rural o urbana, o el nivel socioeconómico de los mismos son variables fundamentales en las finanzas del hogar. Sin embargo, no es posible suponer que la composición será del 100% en ninguno de los estados que conforman ambas variables. Es importante mencionar que, a mayor proporción de hogares con nivel socioeconómico medio, se espera que sea mayor el nivel de educación de sus integrantes y por consiguiente tendrán mayores beneficios. Algo similar sucede con la variable fuente de ingreso, en condiciones realistas será imposible considerar que la fuente de ingresos de la población económicamente activa sea proporcionada al 100% por un empleo formal.

Al examinar las demás variables, identificamos que el manejo del crédito es una variable que puede ser modelada con mayor facilidad y precisión en el corto plazo. Una vez que los jefes del hogar tienen acceso al crédito: consumo, hipoteca, automotriz, concientizarlos acerca de la importancia de no incurrir en morosidad y mantener un buen historial crediticio se logra con educación financiera. Así, *ceteris paribus*, de todas las variables excepto manejo del crédito, se ejecuta el modelo, los resultados muestran que un manejo adecuado del crédito (100%) genera un incremento de 14% en la probabilidad de FH estables respecto al estado base, es decir 80% de probabilidad de FH estables (véase el diagrama 4). Por otro lado, un pésimo manejo del crédito (0%), empeora las FH notablemente; en este caso obtuvimos un incremento en el estado no estable de las FH al pasar de 34% en el estado inicial a 63% (véase el diagrama 5).

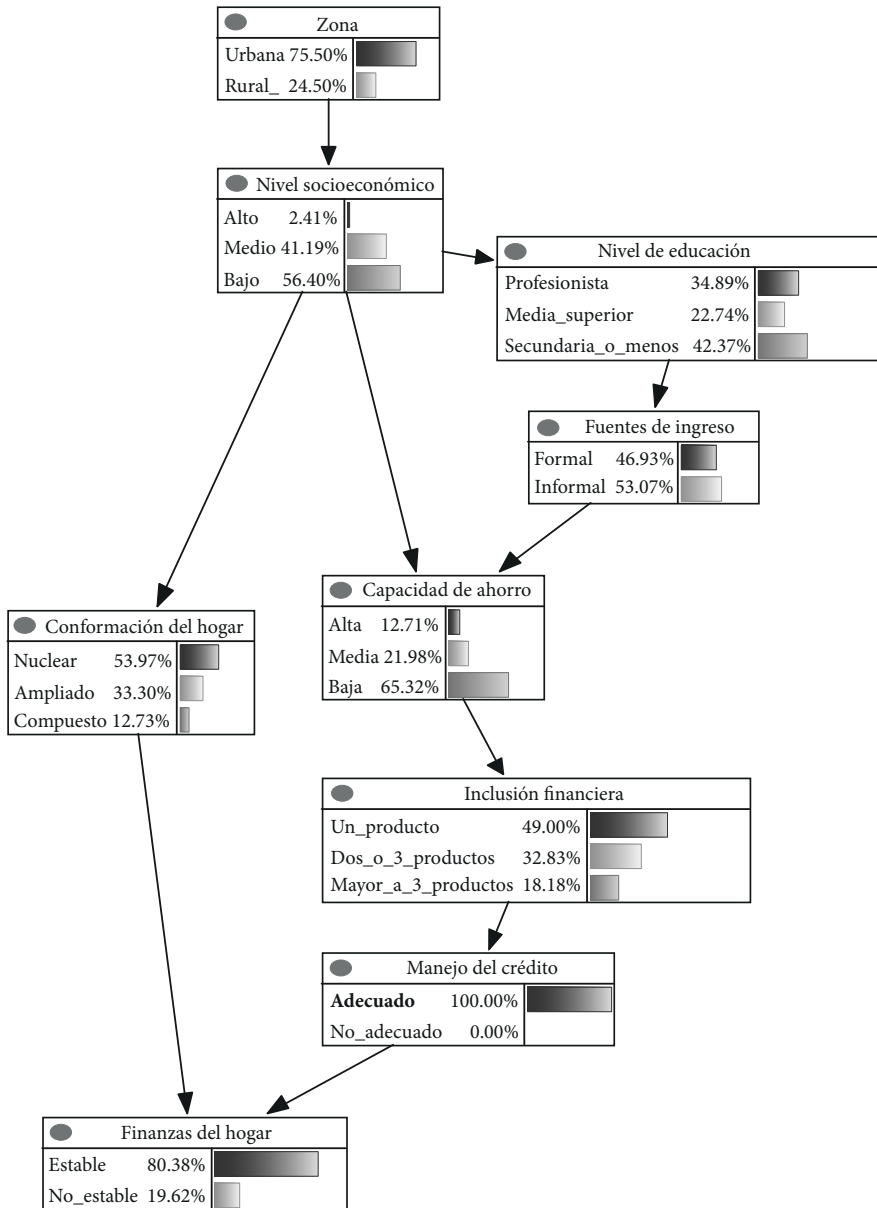
La conformación de los hogares no es una variable en la que se pueda influir directamente. Sin embargo, suponiendo que el 100% de los hogares fuera del tipo compuesto, observamos una disminución de 10.69% en la probabilidad de presentar FH estables (véase el diagrama 6). Sin duda este resultado es relevante, pues revela que más miembros en un hogar no favorece las condiciones financieras del mismo.

Cuando realizamos el análisis de escenarios con las demás variables de la RB, observamos que los impactos a la variable FH no fueron significativos, por lo que no se presentan sus resultados. Sin duda alguna llevar a cabo este ejercicio en dos años más permitirá identificar los efectos a corto plazo de la pandemia de COVID19 sobre el estado de las finanzas de los hogares en México.

5. CONCLUSIONES Y LIMITACIONES

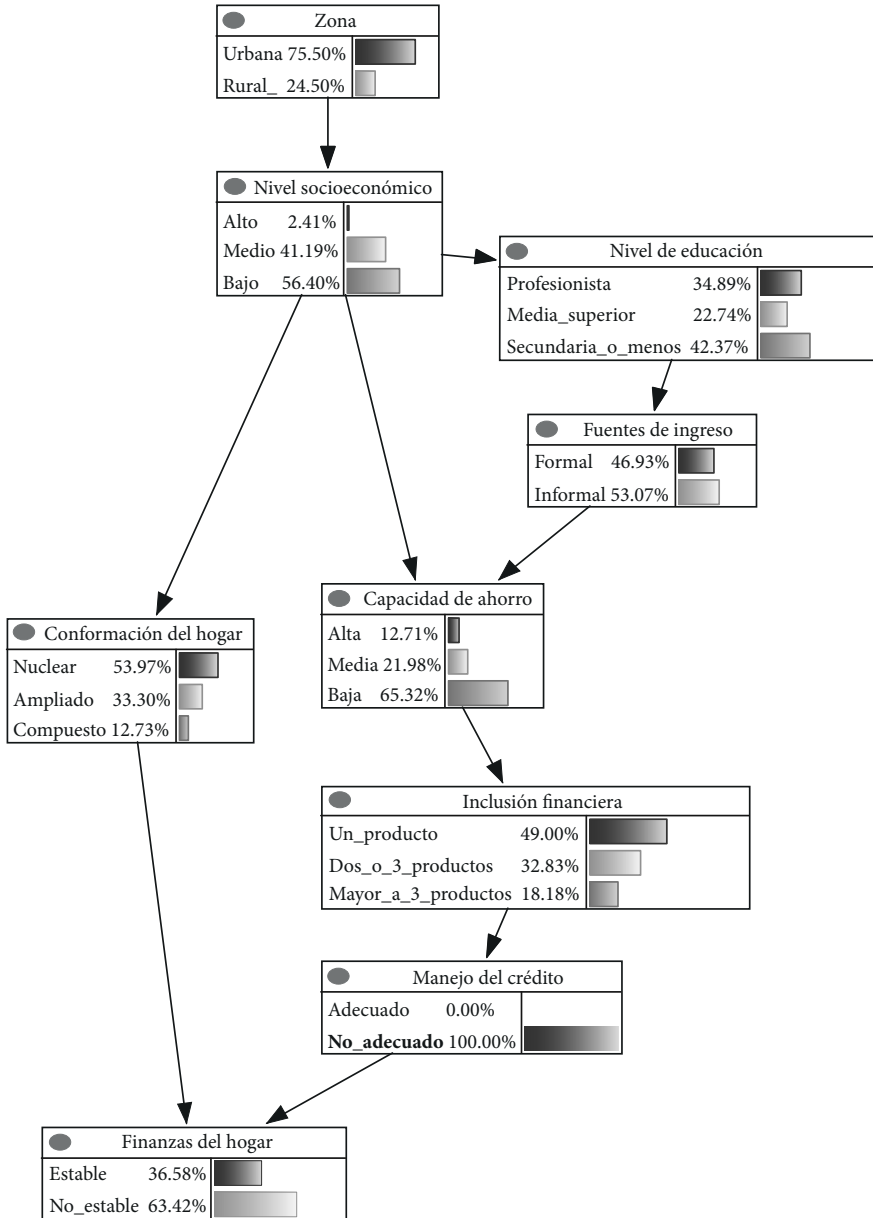
Podemos concluir que el uso de las RB proporciona una forma simple, rápida y eficiente para analizar los diferentes escenarios que se pueden presentar en el estudio del comportamiento de las finanzas de los hogares en México. Gracias a este análisis hemos podido identificar que la variable más sensible para alcanzar FH estables en los hogares mexicanos es el manejo prudente de los créditos que tengan contratados. Un poco menos decisiva, aunque no despreciable es la conformación de los hogares, donde se presenta el hecho de que los hogares compuestos

Diagrama 4. Escenario 1: modelo gráfico del manejo adecuado del crédito



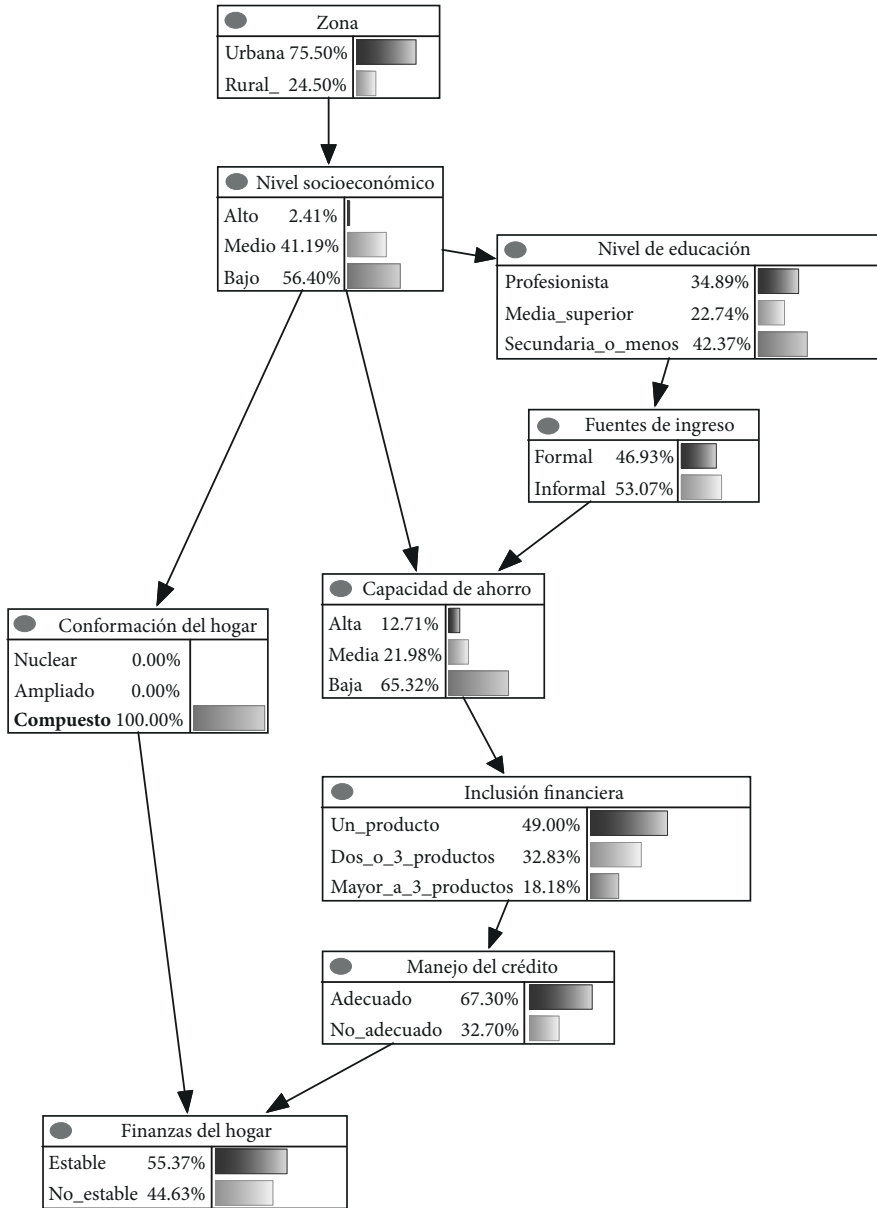
Fuente. elaboración propia empleando el software GeNIe.

Diagrama 5. Escenario 2: modelo gráfico de manejo no adecuado del crédito



Fuente. elaboración propia empleando el software GeNIe.

Diagrama 6. Escenario 3: modelo gráfico de conformación del hogar compuesto



Fuente. elaboración propia empleando el software GeNIe.

no necesariamente mejoran su situación financiera, incluso son más vulnerables.

A partir de los resultados obtenidos, podemos inferir que se requieren promover políticas públicas que se enfoquen en complementar el currículo escolar, que incorporen la impartición de educación financiera en todos los niveles, modalidades y subsistemas educativos; con la intención de destacar la trascendencia que tiene para cada hogar y en especial para aquellos con menores recursos económicos, el manejo adecuado de los créditos y la actividad financiera en general, pues es común que las condiciones crediticias aplicadas a las personas con escasos recursos económicos y académicos, sean las menos favorables. Esto induce perniciosamente a la permanencia de los integrantes de estos hogares en “círculos viciosos” como deudores de prestamistas y establecimientos que venden a crédito.

Como mencionamos en la introducción, una de las limitantes de este tipo de estudios radica en la dificultad de conocer con mayor precisión y detalle la interacción entre las variables que intervienen en las alternativas de inversión y gasto de las familias, así como el nivel de conocimientos financieros de los tomadores de decisión en cada hogar. De hecho, al estimarse que en muchos hogares hay más de una fuente de ingresos proveniente de inversiones no registradas, además de la confidencialidad de la información, es imposible tener datos estadísticos confiables para medir esta variable, por lo que las estimaciones realizadas pueden subestimar la realidad de los ingresos y gastos realizados en los hogares mexicanos. ◀

REFERENCIAS

- Badarinza, C., Campbell, J. y Ramadorai, T. (2016). International comparative household finance. *Annual Review of Economics*, 8(1), pp. 111-144. <https://doi.org/10.1146/annurev-economics-080315-015425>
- Banuls, V.A. y Salmeron, J.L. (2007). A scenario-based assessment model – SBAM. *Technological Forecasting and Social Change*, 74(6), pp. 750-762. <https://doi.org/10.1016/j.techfore.2006.05.015>
- Beltrán, M., Muñoz, A. y Muñoz, A. (2014). Redes bayesianas aplicadas a problemas de credit scoring. Una aplicación práctica. *Cuadernos de Economía*, 37, pp. 73-86. <https://doi.org/10.1016/j.cesjef.2013.07.001>

- Ben-Gal, I. (2007). Bayesian networks. En: F. Ruggeri y R. Kenett, *Encyclopedia of Statistics in Quality & Reliability*. Nueva York: Wiley & Sons.
- Bodie, Z., Merton, R. y Samuelson, W. (1992). Labor supply flexibility and portfolio choice in a life cycle model. *Journal of Economic Dynamics and Control*, 16, pp. 427-449. [https://doi.org/10.1016/0165-1889\(92\)90044-f](https://doi.org/10.1016/0165-1889(92)90044-f)
- Brown, S., Gaino, G. y Taylor, K. (2013). Household debt and attitudes towards risk. *Review of Income and Wealth*, 59, pp. 283-304. <https://doi.org/10.1111/j.1475-4991.2012.00506.x>
- Campbell, J. (2006). Household finance. *The Journal of Finance*, LXI(4), pp. 1553-1604.
- Campbell, J. y Viceira, L. (2002). *Strategic Asset Allocation: Portfolio Choice for Long-term Investors*. Oxford University Press.
- CEEY (2020). Elementos de un Plan Integral para atender las consecuencias económicas de la pandemia de coronavirus en México. [en línea] Disponible en: <<https://ceey.org.mx/elementos-de-un-plan-integral-para-atender-las-consecuencias-economicas-de-la-pandemia-de-coronavirus-en-mexico/>> [Consultado el 22/06/2020].
- Chan, A.P.C., Wong, F.K.W, Hon, C.K.H. y Choi, T.N.Y. (2018). A Bayesian network for reducing accident rates of electrical and mec.hanical (E&M) work. *International Journal of Environmental Research and Public Health*, 15(11), pp. 1-19. <https://doi.org/10.3390/ijerph15112496>
- CONEVAL (2019). *Medición de la pobreza 2008-2018*. Consejo Nacional de Evaluación de la Política de Desarrollo Social (CONEVAL), CDMX. <https://doi.org/10.24275/uami.qj72p713s>
- Cowell, R.G., Dawid, P., Lauritzen, S. y Spiegelhalter, D. (1999). *Probabilistic Networks and Expert Systems*. Nueva York: Springer-Verlag.
- D'Urso, P. y Vitale, V. (2020). Bayesian networks model averaging for best indicators. *Social Indicators Research*, 151(3), pp. 897-919. <https://doi.org/10.1007/s11205-020-02401-z>
- INEGI (2018). *Encuesta Nacional de Ingresos y Gastos de los Hogares en México* (ENIGH). Instituto Nacional de Estadística y Geografía (INEGI). [en línea] Disponible en: <https://www.inegi.org.mx/contenidos/programas/enigh/nc/2018/doc/enigh2018_ns_presentacion_resultados.pdf> [Consultado el 14/06/2020].
- INEGI (2019). *Encuesta Nacional de Ocupación y Empleo*. Instituto Nacional de Estadística y Geografía (INEGI). [en línea] Disponible en: <<https://www.inegi.org.mx/programas/enoe/15ymas/>> [Consultado 12/06/2020].

- INEGI (2020). *Censo Nacional de Población y Vivienda 2020*. Instituto Nacional de Estadística y Geografía (INEGI). [en línea] Disponible en: <<http://www.cuentame.inegi.org.mx/poblacion/hogares.aspx?tema=P>> [Consultado el 10/06/2020].
- Gogolin, F., Dowling, M. y Cummins, M. (2017). Individual values and household finance. *Applied Economics*, 49(35), pp. 3560, 3578. <https://doi.org/10.1080/00036846.2016.1262528>
- Georgarakos, D., Haliassos, M. y Pasini, G. (2014). Household debt and social interactions. *Review of Financial Studies*, 27, pp. 1404-1433. <https://doi.org/10.1093/rfs/hhu014>
- Guidici P. (2004). Integration of qualitative and quantitative operational risk data: A Bayesian approach. En: M.G. Cruz (ed.), *Operational Risk Modeling and Analysis. Theory and Practice* (pp. 131-138), Londres: RISK Books.
- Huber F. (2009). Belief and degrees of belief. En: F. Huber y C. Schmidt-Petri (eds.), *Degrees of Belief. Synthese Library (Studies in Epistemology, Logic, Methodology, and Philosophy of Science)*, Vol. 342 (pp. 1-33). Dordrecht: Springer. https://doi.org/10.1007/978-1-4020-9198-8_1
- Madsen, A.L. y Kjærulff, U.B. (2008). *Bayesian Networks and Influence Diagrams: A Guide to Construction and Analysis*. Nueva York: Springer. <https://doi.org/10.1007/978-0-387-74101-7>
- Mitchell, T.M. (1997). *Machine Learning*. Maidenhead, RU: McGraw-Hill.
- Neil, M., Fenton, M. y Taylor, M. (2005). Using Bayesian networks to model expected and unexpected operational losses. *Risk Analysis*, 25(4), pp. 963-972. <https://doi.org/10.1111/j.1539-6924.2005.00641.x>
- Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems*. San Francisco: Morgan Kaufmann.
- Pearl, J. y Russel, S. (2001). Bayesian networks. Report (R-277), noviembre 2000. En: M. Arbib (ed.), *Handbook of Brain Theory and Neural Networks* (pp. 157-160). Cambridge: The MIT Press.
- Pinheiro, M., Calado P., Silvera M.L., Silva I., Muntz R. y Ribeiro-Neto, B. (2003). Bayesian belief networks for IR. *International Journal of Approximate Reasoning*, 34(2003), pp. 163-179. <https://doi.org/10.1016/j.ijar.2003.07.006>
- Polkovnichenko, V. (2005). Household portfolio diversification: A case for rank-dependent preferences. *Review of Financial Studies*, 18(4), pp. 1467-1502. <https://www.jstor.org/stable/3598027>
- CNBV (2019). *Reporte Nacional de Inclusión Financiera*. 9ª edición. Comisión Nacional Bancaria y de Valores (CNBV). [en línea] Disponible en: <<https://>

- www.gob.mx/cnbv/articulos/9-reporte-nacional-de-inclusion-financiera-rnif9?idiom=es [Consultado el 25/06/2020].
- Spirtes, P., Glymour, C. y Schienens, R. (1993). *Causation Prediction and Search*. Nueva York: Springer-Verlag.
- Supatgiat, C., Kenyon, C. y Heusler, L. (2006). Cause-to-effect operational-risk quantification and management. *Risk Manag*, 8, pp. 16-42. <https://doi.org/10.1057/palgrave.rm.8250001>
- Uusitalo, L. (2007). Advantages and challenges of Bayesian networks in environmental modelling. *Ecological Modelling*, 203(3-4), pp. 312-318. <https://doi.org/10.1016/j.ecolmodel.2006.11.033>
- Weida, E.B., Phojanakong, P., Patel, F. y Chilton, M. (2020). Financial health as a measurable social determinant of health. *PLoS ONE*, 15(5), e0233359. <https://doi.org/10.1371/journal.pone.0233359>