

“EVALUACIÓN DEL PADRÓN DE BENEFICIARIOS DEL SISTEMA DE PROTECCIÓN SOCIAL EN SALUD PARA LA CORRECTA CLASIFICACIÓN DE BENEFICIARIOS”

INFORME EJECUTIVO

México, D.F. a 29 de diciembre de 2014

Este documento presenta los resultados del proyecto sobre “EVALUACIÓN DEL PADRÓN DE BENEFICIARIOS DEL SISTEMA DE PROTECCIÓN SOCIAL EN SALUD PARA LA CORRECTA CLASIFICACIÓN DE BENEFICIARIOS” en el marco de la Licitación Pública LA-012000976-N35-2014 emitida por la *Comisión Nacional de Protección Social en Salud*.

El objetivo general del presente proyecto es realizar un estudio de validación de la Cédula de Características Socioeconómicas del Hogar (CECASOEH) y del modelo estadístico de clasificación utilizado actualmente por el Sistema de Protección Social en Salud.

Los objetivos específicos son:

1. Estimar la pertinencia de la medición obtenida por el instrumento CECASOEH en términos de su sensibilidad y pertinencia, mediante la comparación del instrumento que clasifica a los miembros del padrón del SPSS con un estándar con firmeza conceptual (que se denomina como de oro) que permita la construcción del nuevo régimen contributivo.
2. Revisar el modelo estadístico actual (análisis discriminante) utilizado por la CECASOEH, con el fin de analizar su pertinencia en cuanto a su confiabilidad y validez para mejorar el mismo o bien proponer una alternativa con el fin de que este nuevo tipo de análisis permita definir, con base en evidencia, los deciles socioeconómicos u otra alternativa de medición para la afiliación de familias.
3. Generar un marco conceptual y técnico para proponer el modelo estadístico para el instrumento de CECASOEH, basado en la investigación actual y relevante tanto aplicada como académica sobre los estudios de medición socioeconómica.
4. Proponer un modelo de actualización de las variables que forman parte del instrumento de captación para una CECASOEH versión 2.0 y del algoritmo del modelo estadístico de clasificación de las familias en deciles de ingreso u otro tipo de medida en función de resultados del desarrollo analítico y en campo.

5. Pilotear en campo, la efectividad del nuevo instrumento de captación de información (CECASOEH versión 2.0) o en su defecto, la propuesta de un nuevo instrumento de clasificación de familias en deciles de ingreso o la alternativa de medición.
6. Considerar que el modelo de medición y el instrumento depende de los procesos logísticos de incorporación de familias realizados en los Módulos de Afiliación y Orientación al SPSS.
7. Revisar el proceso de incorporación de familias al SPSS, con énfasis en el proceso de registro de información, con el fin de identificar áreas susceptibles de mejora.

Específicamente, se analizan tres dimensiones subyacentes al modelo de selección utilizado por la SPSS para clasificar a la población beneficiaria de acuerdo a los distintos esquemas de aportación (cuotas):

- a) Las variables a incorporar;
- b) La ponderación de cada dimensión;
- c) El umbral a partir del cual se considera que un hogar o individuo presenta la carencia de cada dimensión de bienestar

Descripción de la metodología

La identificación de las variables a incorporar así como su ponderación en el proceso de identificación de beneficiarios, se llevó a cabo por medio de análisis discriminante que permite identificar los espacios que definen dos condiciones, la de pertenecer al grupo de potencial beneficiario de un programa social, y la de no serlo, a partir de arreglos multidimensionales que caracterizan estas condiciones, mediante valores típicos de las variables utilizadas. Partiendo de la información del ingreso y comparándolo con un punto de referencia establecido, se dispone de una base para caracterizar los perfiles que distinguen a los hogares con mayor y menor bienestar mediante un conjunto de indicadores.

La metodología de análisis discriminante es una herramienta que cuenta con los elementos y la robustez para constituirse como mecanismo de focalización para el SPSS y por eso se contempla como principal alternativa metodológica. En el caso del proyecto actual, el contar con un mecanismo de focalización es esencial para evitar errores de inclusión (es decir, catalogar a hogares que no pertenecen a los primeros cuatro deciles de medición del bienestar) o de exclusión (que consiste en clasificar erróneamente a hogares que pertenecen a los primeros cuatro deciles de

medición del bienestar, como si pertenecieran a grupos de mayor nivel de bienestar en deciles superiores).

La importancia de evitar errores de inclusión y exclusión en el marco del SPSS radica en que un error de inclusión eximiría del pago del servicio a hogares que están en condición de realizar alguna aportación, mientras que el error de exclusión llevaría a exigir una cuota de pago a hogares a los que no les correspondería realizar una aportación debido a que cuentan con las condiciones para clasificarse (bajo una metodología con mayor precisión) a los primeros cuatro deciles de nivel de bienestar.

Principales resultados y conclusiones

Las conclusiones del análisis del modelo utilizado actualmente por la SPSS para la clasificación de beneficiarios en la distribución del ingreso –lo cual se estima por medio de la aplicación del algoritmo empleado a los datos de la Encuesta Nacional de Ingreso Gasto de los Hogares (ENIGH) 2012 que la fuente más fidedigna con la que se cuenta para obtener información sobre el ingreso de toda la distribución de población del país- son que:

- En términos de las variables, la CECASOEH brinda un punto de partida correcto al considerar variables de distintas dimensiones que se complementan y logran caracterizar el estrato socio-económico al no contar con indicadores directos viables de ingreso.
- El modelo de clasificación actual del SPSS asigna correctamente al 14.4% de los hogares de acuerdo a su posición real en la distribución del ingreso.
- En general el modelo es un mejor predictor para los deciles bajos (del 1 al 4), reduciendo su precisión para el resto de los deciles, particularmente para el 9 y el 10, donde el margen de error es cercano a cero.
- Considerando estratos rural y urbano por separado, en lo rural el modelo predice con un alto grado de precisión para los primeros cuatro deciles rurales, pero un bajo nivel de precisión para el resto. Por otra parte, para lo urbano, el modelo es un buen predictor para los primeros cuatro deciles y los dos últimos, lo cual puede servir como referente.

Utilizando las mismas variables del modelo de clasificación del SPSS actual, se aplicó la metodología de análisis discriminante sobre los datos de la ENIGH para evaluar si mejora o no el grado de predicción en la clasificación respecto a la distribución por deciles de la ENIGH. La utilización de esta metodología implica el reestimar los ponderadores para identificar los que sean más idóneos para incrementar el poder de precisión del instrumento lo más posible. Los resultados dicha aplicación son:

- Cuando las variables originalmente incorporadas en la CECASOEH se procesan utilizando el análisis discriminante el poder de predicción de dichas variables se incrementa considerablemente. Con este modelo, se asigna atinadamente al 47.4% de los hogares, mejorando así el grado de precisión del modelo actual de clasificación del SPSS, que ubica adecuadamente solo al 14.4%.

Como resultado del análisis llevado a cabo en el presente estudio, se identifica un conjunto de variables que mantiene como base las variables planteadas por el modelo actual del SPSS complementándolas con un conjunto de indicadores que cuentan con mayor precisión para la captación de tenencia de activos y mejor caracterización de los deciles más altos. Las variables para el sector rural en dicha propuesta son las siguientes:

- Material del piso de la vivienda
- Si cuenta con excusado
- Educación del jefe del hogar
- Índice de dependencia
- Hacinamiento
- Seguro social
- Hijos que no asisten a la escuela
- Hijos que trabajan
- Tipo de servicios de salud
- Si cuenta con estufa
- Si cuenta con refrigerador
- Si cuenta con lavadora
- Número de prestaciones de todos los que trabajan en el hogar
- Si cuenta con vehículo 0, 1 o más de 2

Las variables urbanas son las siguientes:

- Tamaño de localidad
- Si cuenta con excusado
- Educación del jefe del hogar
- Índice de dependencia
- Hacinamiento
- Seguro social
- Hijos que no asisten a la escuela
- Hijos que trabajan
- Si cuenta con vehículo 0, 1 o más de 2

- Si cuenta con computadora 0, 1 o más de 2
- Número de prestaciones de todos los que trabajan en el hogar
- Número de integrantes que asisten a escuela pública
- Número de personas que laboran en el hogar entre el total de integrantes
- Tipo de empresa en la que labora (1. Tipo independiente, familiar o personal)
- Tenencia de la vivienda
- Número de miembros a los que les otorgaron beca escolar
- Si cuenta con celular
- Si cuenta con teléfono
- Tipo de servicios de salud
- Si cuenta con TV de paga
- Si cuenta con conexión a internet

Los resultados de la aplicación de este modelo, denominado CECASOEH v 2.0, utilizando análisis discriminante, y estimando su poder de decisión por medio de la ENIGH, son:

- Con la agregación en cuatro grupos y la combinación de modelos, se logra incrementar el poder de precisión a 56.6% de los hogares. Es decir, la metodología predice adecuadamente la ubicación de 56.6% de los hogares en el grupo socioeconómico que realmente le corresponde.
- Sin embargo, si se hace la agregación considerando únicamente tres grupos de deciles, el modelo combinado arroja como resultado propuesto permite mantener el alto grado de predicción para el primer grupo de deciles (1 al 4) y mejorar significativamente el grado de predicción para el grupo de deciles intermedios (5 al 8) y para el grupo de últimos deciles (9 y 10). Con este modelo, se clasifica atinadamente al 66.36% de los hogares, siendo este el grado de predicción más alto entre los ejercicios realizados.

Los análisis aquí planteados y su comparación permiten concluir lo siguiente:

- Es necesario evaluar, desde la perspectiva práctica de la implementación del sistema, los objetivos que éste busca y en base al grado de precisión en la clasificación de los distintos grupos de deciles, si se requiere modificar los límites y condiciones de pertenencia para los regímenes contributivo y no contributivo. Por ejemplo, una alternativa es considerar un modelo que identifique con la mayor precisión posible a los últimos deciles de ingreso y establecer el régimen contributivo sólo para estos. Lo anterior requiere tomar en cuenta

todas las consideraciones relevantes en cuanto al diseño del SPSS como política pública, la viabilidad operativa y el objetivo implícito de contar con la participación del afiliado en el co-financiamiento de su aseguramiento en salud con las consecuencias que esto conlleva.

- Dado el alto grado de predicción que el modelo actual del SPSS muestra para el primer grupo de deciles, aunado al incremento en el grado de precisión para los otros dos grupos de deciles (particularmente los últimos deciles) producto del planteamiento de un modelo alternativo, la propuesta de este estudio radica en combinar ambas alternativas en un modelo que, comparado con el resto de los ejercicios realizados, permite asignar a un mayor porcentaje de los hogares atinadamente; o en otras palabras, muestra el menor margen de error.
- El cuadro siguiente resume los distintos modelos contemplados en el Análisis

Cuadro Resumen:

Comparación de análisis/ejercicios realizados respecto al % de hogares asignados atinadamente

Análisis/Ejercicio	% de hogares asignados atinadamente
Modelo de clasificación del SPSS actual, agregado en deciles	14.40%
Modelo de clasificación del SPSS actual, agregado en cuatro grupos de deciles	41.25%
Modelo de análisis discriminante con variables del modelo del SPSS actual, agregado en cuatro grupos de deciles	47.37%
Modelo combinado: algoritmo rural del modelo del SPSS actual + modelo de análisis discriminante con variables nuevas, agregado en cuatro grupos de deciles	56.64%
Modelo combinado: algoritmo rural del modelo del SPSS actual + modelo de análisis discriminante con variables nuevas, agregado en tres grupos de deciles	66.36%

Fuente: Elaboración propia con resultados de ejercicios realizados

- Asimismo, la clasificación en tres grupos o estratos coincide con la lógica actual de separación de regímenes y asignación adicional a estos por consideraciones particulares

(contar con un menor de 5 años en el hogar o con una embarazada o ser beneficiarios de Prospera, entre otros).

En conclusión, se propone valorar la posibilidad de utilizar este modelo combinado alternativo con tres grupos de deciles o de clasificación como la mejor opción para lograr los objetivos propuestos.

Análisis de viabilidad de la propuesta por medio del piloteo de la CECASOEH v 2.0 en campo

Como resultado del análisis, investigación y recopilación de información, se diseñó una propuesta del instrumento de clasificación socioeconómica CECASOEH v 2.0 que incorpora las variables mencionadas en el apartado anterior.

El muestreo para la aplicación del instrumento se realizó mediante la selección de localidades representando estratos en zona urbana (78%) y rural (22%) de acuerdo con el Censo 2010 de INEGI. La muestra estará integrada por 3,000 hogares a nivel nacional. El grado de error para una muestra de tamaño 3,000 viviendas al 95% de confianza es de +1.79% considerando un procedimiento de muestreo aleatorio simple.

El operativo de campo se realizó en la Zona Metropolitana de la Ciudad de México, Jalapa, Tuxtla Gutiérrez, Puebla, León, Oaxaca. En total, las cuales cumplen con el criterio de selección y distribución urbano/rural, y conforman una muestra con representatividad de distintos niveles socioeconómicos. El levantamiento se realizó en la ciudad y la zona rural cercana a estas ciudades. El Público Objetivo fueron Hombres y Mujeres mayores de edad del hogar seleccionado.

Durante el levantamiento se aplicó la nueva versión de la Cédula de Características Socioeconómicas del Hogar (CECASOEH v 2.0) considerando las variables analizadas en la sección anterior. El levantamiento de campo se realizó del 06 al 19 de diciembre del 2014. Para lograr las 3000 entrevistas efectivas se realizaron 11,999 contactos, es decir el 25% de los casos se concretaron con éxito y cumplieron con los criterios de calidad. A continuación se especifica el tipo de incidencias registradas.

Tipo de Incidencias

TIPO DE INCIDENCIAS	Total	%
Entrevistas efectivas y validadas	3000	25.0%

No desea participar	2536	21.1%
Entrevista cortada / cancelada	301	2.5%
Contacto con menor de edad / no respondiente adecuado	590	4.9%
Hogar deshabitado / abandonado	948	7.9%
Predio vacío	867	7.2%
Sin acceso / Restricción para ingreso	590	4.9%
No abrió	3061	25.5%
Otro	106	0.8%
TOTAL	11,999	100.0%

A continuación se describe cada uno de los tipos de incidencias encontradas en campo:

- **Entrevistas efectivas y validadas.** Cuando se obtiene información de todos los miembros de hogar y de todas las secciones del cuestionario de manera satisfactoria.
- **No desea participar.** Cuando se realiza algún contacto con el posible entrevistado y éste expresa que no le interesa o no quiere participar.
- **Entrevista cortada/ cancelada.** Cuando el entrevistado, una vez aceptada la entrevista, decide ya no continuar por alguna razón. Cabe mencionar que este tipo de incidencias representan menos del 2,5% del total de contactos realizados.
- **Contacto menor de edad / no es el respondiente adecuado.** Cuando la persona que se encuentra en el hogar aún no cumple la mayoría de edad y por la tanto no es candidato para la entrevista.
- **Hogar deshabitado / abandonado.** Se clasifican las viviendas aparentemente ocupadas pero que al momento de tocar no salió nadie. Se indaga con los vecinos e informan que en este momento nadie la habitaba.
- **Predio vacío.** Cuando de acuerdo al salto sistemático, el hogar que se debía tocar es un lote vacío.
- **Sin acceso / Restricción para ingreso.** Cuando el hogar se encuentra en alguna cerrada, edificio, etc. y que no se puede tener un contacto directo con el posible entrevistado.
- **No abrió.** Cuando después de insistir nadie abrió en el hogar.
- **Otros.** Cualquier caso no cubierto en los anteriores. Por ejemplo:
 - Se encontraron casos en los que el acceso a la zona de trabajo fue una barrera debido a que ésta era una zona considerada de alto riesgo.
 - En zonas rurales no se encontraron datos de identificación del lugar ni nomenclaturas o vías de acceso.
 - Se mostró resistencia por parte de las autoridades locales quienes retuvieron a miembros del equipo de campo para indagar sobre la actividad realizada así como el fin de la misma. En ningún caso se registraron incidentes graves.

En general no se observa mayor dificultad para la aplicación del instrumento, únicamente en 301 casos de todos los contactos realizados (2.5%) el informante decidió dar por terminada la entrevista sin completar toda la información del cuestionario.

Como marco de referencia, en un levantamiento de campo regular observamos una tasa de abandono promedio del 30% durante los primeros 10 minutos de aplicación, esto quiere decir que 3 de cada 10 entrevistas que comienzan se dan por terminadas antes de que ésta sea finalizada completamente.

Con lo anterior se puede deducir que el instrumento se encuentra en una tasa de abandono muy por debajo de lo normal, lo que pronostica un desempeño muy positivo en términos de esfuerzo para el personal de campo, así como interés por parte del entrevistado y poca percepción de preguntas intrusivas.

Las dos principales secciones donde los entrevistados mostraron mayor renuencia a dar información fueron:

- Datos personales de los integrantes del hogar y parentesco con el jefe del hogar.
- Artículos con los que cuenta en su hogar.

En ambos casos, los informantes adjudicaron su abandono a razones de seguridad y falta de confianza, la labor del entrevistador para transmitir confianza y certidumbre a los respondentes acerca del operativo permitió que el cuestionario se completara hasta darlo por terminado.

Sin embargo, no se prevé que dichas razones afecten un proceso real de incorporación a los servicios ofrecidos por la SPSS debido a que la afiliación es voluntaria y requiere de la identificación plena de los miembros del hogar. Por este motivo, la conclusión es que el instrumento es adecuado para el propósito establecido.

En cuanto a la implementación de las entrevistas se registra un tiempo promedio de aplicación de 08:29 minutos desde la aceptación a participar en la entrevista por parte del respondente y hasta la terminación del registro de todos los datos solicitados.

- Se registra una desviación estándar de .00190 minutos.
- El 67% de las entrevistas han durado entre 6 y 9 minutos.