

# Estimaciones, identidad y la relación entre uso de la lengua y autoadscripción

## Estimates, identity and relationship between language use and self-ascription

Ignacio Ibarra-López, Héctor Calleros y Jorge David Cortés-Moreno

*Tecnológico de Monterrey, México; Centro de Estudios Americanos, Universidad de Varsovia, Polonia; Instituto de Ciencias de Gobierno y Desarrollo Estratégico de la Benemérita Universidad Autónoma de Puebla, México*

### Resumen

Este trabajo, estudia las diferencias que surgen cuando se utiliza la autoadscripción o el uso de la lengua indígena en la estimación de la población indígena. Producto del análisis de la información oficial, se concluye que el aumento en la escolaridad en el hogar y el recibir ingresos de familiares que viven en el extranjero, desincentiva a las personas a indicar el uso de lengua y autoadscribirse. Por otro lado, se encuentra que las personas que reciben apoyos por parte del gobierno tienen mayores incentivos a declarar el uso de la lengua y autoadscribirse. Finalmente, se demuestra que autoadscripción y uso de la lengua indígena no son las mismas variables y por lo tanto no pueden utilizarse como un criterio equivalente para la realización de la política pública. Se demuestra estadísticamente que las personas que hablan la lengua indígena constituyen un subconjunto de las personas que se autoadscriben como indígenas.

*Palabras clave:* Población indígena, autoadscripción, censos, comportamiento estratégico.

### Abstract

This work studies the differences that arise when self-identification or the use of the indigenous language is used in the estimation of the indigenous population. As a result of the analysis of official information, it is concluded that the increase in schooling at home and the receipt of income from relatives living abroad, discourages people from indicating the use of the language and self-registering. On the other hand, it is found that people who receive support from the government have greater incentives to declare the use of the language and self-enroll. Finally, it is shown that self-enrollment and use of the indigenous language are not the same variables and therefore cannot be used as an equivalent criterion for the realization of public policy. It is statistically demonstrated that people who speak the indigenous language constitute a subset of people who self-describe as indigenous.

*Key words:* Indigenous population, self-identification, censuses, strategic behavior.

## ANTECEDENTES

**D**isposiciones a nivel nacional y a nivel internacional establecen que los pueblos originarios tienen el derecho individual y colectivo de mantener el desarrollo de sus propias características e identidades, así como reconocer a sus miembros y ser reconocidos de esa forma.<sup>1</sup> Para ello, se ha determinado el criterio de autoadscripción. Si una persona o grupo de personas se auto-adscriben como descendiente de un pueblo originario en una comunidad, es suficiente para considerar que existe un vínculo cultural, histórico, lingüístico o de otro tipo con dicha comunidad y eso le permite ser gobernada por regulaciones especiales (e.g. usos y costumbres).<sup>2</sup>

México ratifica el Convenio 169 (C169) de la Organización Internacional del Trabajo (OIT) en el año de 1990, según Decreto publicado en el Diario Oficial de la Federación.<sup>3</sup> En este documento, se indica que un pueblo indígena o tribal, constituye todas las personas que descienden de poblaciones que habitaban en la época de la conquista o la colonización y que conservan todas sus propias instituciones sociales, económicas, culturales y políticas, o parte de ellas (ver artículo 1.1).<sup>4</sup> Adicionalmente, en el artículo 1.2, establece que: "...la conciencia de su identidad indígena o tribal deberá considerarse un criterio fundamental para determinar quién se debe considerar indígena en una sociedad".

La Corte Interamericana de Derechos Humanos (CIDH) ha realizado un análisis<sup>5</sup> del cual se puede destacar lo siguiente (CIDH, 2009):<sup>6</sup> i) los elementos que definen a un pueblo indígena son tanto objetivos como subjetivos, donde el elemento subjetivo corresponde a la autoidentificación colectiva en tanto pueblo indígena;<sup>7</sup> ii) los elementos que permiten identi-

<sup>1</sup> A nivel nacional véase Artículo 2, quinto párrafo de la Constitución Política de los Estados Unidos Mexicanos; a nivel internacional véase Artículo 1 Sección 2 del Convenio Número 169 de la Organización Internacional del Trabajo y los Artículos 3,4,9 y 32 de la Declaración de las Naciones Unidas con respecto a los Derechos de las Poblaciones Indígenas. Se recomienda revisar la ratificación de México a este convenio disponible en: [http://www.ilo.org/dyn/normlex/es/f?p=NORMLEXPUB:12100:0::NO:12100:P12100\\_INSTRUMENT\\_ID:312314:NO](http://www.ilo.org/dyn/normlex/es/f?p=NORMLEXPUB:12100:0::NO:12100:P12100_INSTRUMENT_ID:312314:NO).

<sup>2</sup> Véase: [http://portal.te.gob.mx/en/contenido/\\_indigenous\\_-communities-\\_self-adscription-criteria-e](http://portal.te.gob.mx/en/contenido/_indigenous_-communities-_self-adscription-criteria-e)

<sup>3</sup> Véase: [http://www.cdi.gob.mx/transparencia/convenio169\\_oit.pdf](http://www.cdi.gob.mx/transparencia/convenio169_oit.pdf)

<sup>4</sup> Véase: [http://www.ilo.org/\\_dyn/normlex/es/f?p=NORMLEXPUB:12100:0::NO:12100:P12100\\_INSTRUMENT\\_ID:312314:NO](http://www.ilo.org/_dyn/normlex/es/f?p=NORMLEXPUB:12100:0::NO:12100:P12100_INSTRUMENT_ID:312314:NO)

<sup>5</sup> Véase: <http://cidh.org/countryrep/TierrasIndigenas2009/Cap.III-IV.htm>

<sup>6</sup> Véase: <http://www.oas.org/es/cidh/indigenas/docs/pdf/tierras-ancestrales.esp.pdf>

<sup>7</sup> La CIDH retoma el documento de la OIT, "Los Derechos de los Pueblos Indígenas y Tribales en la Práctica - Una Guía sobre el Convenio N°. 169 de la OIT". Programa para promover el Convenio Núm. 169 de la OIT (PRO-169), Departamento de Normas Internacionales del Trabajo, 2009, pág. 9.

ficar a una persona perteneciente a un pueblo indígena no son una definición inclusiva o comprehensiva sino factores que pueden estar presentes en mayor o menor grado en distintas regiones y contextos nacionales y locales;<sup>8</sup> iii) al establecer los Derechos de los Pueblos Indígenas, no se busca definir a los pueblos indígenas por parte del estado, pero si a establecer que los pueblos indígenas tienen derecho a determinar su propia identidad o pertenencia conforme a sus costumbres y tradiciones de forma colectiva o individual (en consonancia con el artículo 33.1, C169).

El reconocimiento de la diversidad cultural está ligado a la política social, tal cual lo han planteado instituciones internacionales como el Banco Interamericano de Desarrollo (BID). Según este organismo, una de las principales estrategias que se debe perseguir, es la de visibilizar a los pueblos indígenas, incorporando sus temas en las agendas de desarrollo en los países. Para lograr este objetivo, se debe garantizar que exista una medición adecuada de la población indígena en censos, conteos y encuestas de hogares, incorporando las concepciones indígenas en diferentes temas como pobreza, bienestar, procesos participativos, así como en la propia validación de los datos estadísticos.<sup>9</sup>

En México, para identificar a la población considerada descendiente de un pueblo indígena, se utiliza el criterio de uso de la lengua. Se asume que esta variable es capaz de sintetizar la enorme diversidad étnica y pluricultural de México. Desde la óptica del gobierno, este criterio resuelve en lo operativo, el análisis, formulación y evaluación de políticas sociales atendiendo a lo suscrito en el C169 de la OIT.

El siguiente trabajo discute sobre la relación entre el uso de la lengua y la autoidentificación. Se demuestra estadísticamente que existe un sesgo sistemático entre las dos variables. También se exploran algunas de las razones por las cuales quienes se autoadscriben o indican el uso de la lengua, siguen un comportamiento estratégico. En particular, se analiza el impacto de la escolaridad, recibir apoyos del gobierno y recibir ingresos de familiares que viven en el extranjero.

<sup>8</sup> Véase: ONU – Consejo Económico y Social – Comisión de Derechos Humanos – Subcomisión sobre la Prevención de la Discriminación y la Protección de las Minorías – Grupo de Trabajo sobre Poblaciones Indígenas: “Working Paper by the Chairperson-Rapporteur, Mrs. Erica-Irene A. Daes, on the concept of ‘indigenous people’”. Documento ONU E/CN.4/Sub.2/AC.4/1996/2, 10 de junio de 1996, párrs. 69-70.

<sup>9</sup> Ver el documento: Política Operativa Sobre Pueblos Indígenas y Estrategia para el Desarrollo Indígena (BID, 2006). En particular en la sección VI. Ejes estratégicos y prioridades para la acción del Banco. Disponible en: <http://idbdocs.iadb.org/wsdocs/getdocument.aspx?docnum=2032314>

## **CRITERIOS ALTERNATIVOS AL USO DE LA LENGUA PARA LA IDENTIFICACIÓN DE LA POBLACIÓN INDÍGENA**

Según Schkolnik y Del Popolo (2005: 109), en el concepto “población indígena” no debe existir distinción del pueblo al que las personas pertenecen o con el cual se identifican. En opinión de estos autores, esto no ha sido posible porque, aunque se tengan las variables o indicadores en las encuestas oficiales, en realidad no se ha generado una integración de dichas variables para construir una definición tal cual se establece en los propios instrumentos legales internacionales (Schkolnik y Del Popolo: 2005: 112).

En su análisis, descubren que de acuerdo con los criterios establecidos por el uso de la lengua y la autoadscripción, se generan estimaciones diferentes de población. Por lo tanto, que las variables no son intercambiables o equivalentes. Demuestran que sistemáticamente la estimación de la población que retoma el uso de la lengua siempre es superada por la estimación que utiliza la autopertenencia. También concluyen que dentro de la población indígena existen subcategorías que indican que un pueblo indígena no debe ser considerado como un conjunto homogéneo de personas.

En años más recientes, Urrea-Giraldo y Rodríguez-Sánchez (2014) encontraron que existe una disyuntiva importante entre destacar o no las diferencias entre grupos poblacionales. Muestran que una de las consecuencias de incluir en los censos preguntas relacionadas con la cultura o la pertenencia étnica, incrementan con el paso del tiempo, el peso que pueden tener determinados grupos en una sociedad. Por lo tanto, la inclusión o no inclusión de variables que miden la diversidad es contingente a las posturas de los gobiernos<sup>10</sup>. Sostienen que incluir variables y metodologías basadas en diversidad étnica y cultural, es una herramienta para establecer procesos redistributivos del ingreso, reparación del daño o bien acciones afirmativas.

Para el caso de México, también existe una discusión sobre cuál es la forma más efectiva de identificar a la población indígena (Gamio, 1942; Cifuentes y Moctezuma, 2006; Rubio, 2014; Saldívar y Walsh, 2014). Se han propuesto diferentes enfoques con resultados muy diversos. Por ejemplo, se ha observado que al utilizar variables relacionadas con las costumbres de las personas (*e.g.* el vestido, calzado o la alimentación) se genera

<sup>10</sup> En México, Colombia y Perú, se busca una integración nacional por lo cual existe la motivación a no generar distinciones por el color, pertenencia a un grupo étnico o clase social y por ello se han tardado más en incorporar variables que destacan diferencias. En el caso de Brasil ocurre exactamente lo contrario, pues la inclusión de variables que miden la diversidad cultural juega a favor del establecimiento de políticas en contra de la discriminación (Urrea-Giraldo y Rodríguez-Sánchez, 2014: 29-33).

una mayor población indígena (Valdés, 1995; Pla, 2011). Otro criterio que se ha discutido es el del color de piel, lo cual ha sido criticado principalmente por las connotaciones discriminatorias que conlleva, pero también por una cuestión más práctica: dado el mestizaje, es imposible identificar a los indígenas por el tono de piel (Wade, 2008; Knight, 2010; Molina, 2010; Landázuri, 2012).

Schmelkes (2013) propone dos características básicas que pueden identificar a la población indígena de México: el alto grado de marginación y el de mantener instituciones sociales, económicas y culturales propias. Esta autora indica además que el problema de la identificación se propicia por forzar una metodología de conteo de personas, sobre un concepto que en realidad es una construcción jurídica.

Al identificar a las personas indígenas, se enfrentan algunos problemas metodológicos parecidos a los que surgen al identificar a las personas en pobreza. Por esta razón, la identificación de indígenas se asocia (al igual que la pobreza), con cambios políticos que propician elegir cierto tipo de herramientas estadísticas para generar las mediciones (Cifuentes y Moctezuma, 2006).<sup>11</sup>

La variable cultural que se ha discutido con mayor insistencia recientemente en la literatura para el caso de México es la autoadscripción. Esta variable (o al menos una de sus variantes) apareció en el Censo de 1921 (Valdés, 2001), fue descartada en versiones posteriores y nuevamente se incluyó desde el Censo de 2000 a la fecha. Con toda honestidad, la identificación centrada en la autoadscripción también enfrenta una crítica.

Por ejemplo, Perraudin (2010), ha encontrado que algunas personas pueden vestir con ropa tradicional para recibir apoyos del gobierno, identificándose como indígenas. Esto ocurre porque presentarse como indígena puede facilitar el ser candidato a ciertos programas (al igual que identificarse como pobre). Pero para ello, la persona debe tener el dominio del estereotipo del “indígena”, lo cual pudiera “reforzar” su identidad (desde el cliché de ser indígena).

Si bien existe esta connotación negativa de la autoadscripción, también existen posturas a favor de la autoadscripción porque los censos en México se realizan con base en los criterios y necesidades de quienes los

<sup>11</sup> Un ejemplo muy claro es la propia evolución de los censos en México (Rubio, 2014). En 1921 se tenía la pregunta donde explícitamente se le pedía a la persona indicar a que raza pertenecía. En 1940 y 1950, se preguntó por el uso de zapatos, huaraches, si vestía “indumentaria tipo indígena”, por las costumbres alimenticias (e.g. comer trigo o maíz) e incluso por el tipo de lugar donde se dormía (tapesco, catre, cama, hamaca o el suelo).

producen y no de la propia población indígena (Carrasco y Alcázar, 2009). Schmelkes celebra la inclusión de la autoadscripción al decir:

...sería muy conveniente, como ya indicábamos, añadir mayores aproximaciones a la definición de la condición de ser indígena. Celebramos enormemente que se considere ya desde hace 10 años, en el cuestionario ampliado, la autoadscripción, así como que se haya construido la variable de la pertenencia a un hogar donde el jefe del mismo hable una lengua indígena (Schmelkes, 2013).

El tema de identificación de la población indígena a través de indicadores culturales es tan complicado, que incluso existen autores que han declarado que identificar a la población indígena de esta forma, es una tarea que debe sortear diferentes obstáculos que incluyen los procesos de mestizaje y transculturación (Knight, 2010), la discriminación (Janssen y Martínez, 2006) o la negativa de las personas a asumir su identidad frente a encuestadores estatales que pudieran asociar con un “sistema opresivo” (Stavengen, 1992: 67).

El reconocimiento de una comunidad indígena ha sido y es una preocupación en el ámbito académico, debido a la enorme complejidad que implica el proceso de identificación de grupos poblacionales heterogéneos. Existe una controversia por incluir o no incluir variables que destaquen las diferencias culturales, debido a posibles implicaciones políticas y sociales. La autoadscripción es la alternativa que existe en la actualidad, sin que ello implique que esté exenta de críticas, que pueda ser mejorada y complementada por otras variables. A continuación, se discute precisamente la óptica de la autoridad en torno a la autoadscripción.

## **LA VISIÓN DE LA AUTORIDAD EN RELACIÓN CON LA AUTOADSCRIPCIÓN**

Los esfuerzos del gobierno mexicano para identificar a la población indígena inician en 1921, cuando en el Censo se incluyó una pregunta donde las personas se tenían que autoadscribir de acuerdo con su raza. El dato que se obtuvo es que 25 por ciento de los encuestados, se autoadscriben como de “raza indígena” (Pla, 2011). La reacción del gobierno en ese tiempo fue la de eliminar la pregunta, pero se amplió la información sobre el uso de la lengua, distinguiendo entre monolingües y bilingües. Esta decisión inicia la tradición de la identificación a partir del uso de la lengua. El ejercicio de autoadscripción hasta donde se conoce no se retoma sino a partir de que México ratifica el C169 y se establece una muestra con datos del Censo de 2000.

Resulta interesante que el propio INEGI reconoce que el criterio de la lengua presenta algunas limitaciones. En particular que la reducción en el número de lenguas autóctonas no se debe asociar con la reducción de la población indígena (INEGI, 2000). Por ello al retomar el criterio de autoadscripción, se muestra una preocupación de generar estadísticas confiables debido a que puede existir una subdeclaración de las personas indígenas, ligada a prejuicios de índole negativa presentes principalmente en las ciudades (INEGI, 2004).

Debido a que existen discrepancias entre autoadscripción y la estimación que motivaron que la dependencia encargada de los asuntos indígenas<sup>12</sup> realizara en conjunto con el Programa de las Naciones Unidas para el Desarrollo (PNUD) y el Consejo Nacional de Población (CONAPO) estimaciones mediante un enfoque alternativo (Rubio, 2014). Dicha metodología, utilizó la pertenencia a un grupo indígena, uso de la lengua, adscripción y variables relacionadas con características de los hogares. Este estudio encontró que existe una cantidad mayor de personas indígenas.

La postura oficial tras los resultados de estos trabajos fue la de defender el criterio lingüístico, aunque se reconoció que este criterio puede resultar insuficiente como marcador de identidad (CDI-PNUD, 2010: 24). Si bien, los estudios confirmaban la necesidad de integrar de variables alternativas como la autoadscripción, la autoridad dejó de lado esta posibilidad.<sup>13</sup> Esta decisión tiene como consecuencia la imposibilidad de garantizar el reconocimiento en todos los ámbitos de la diversidad étnica (González-Ulloa, 2013; CONAPRED, 2010; 2011; 2012).

Posteriormente, organismos autónomos como el Consejo para la Evaluación de la Política en Desarrollo Social (CONEVAL), reconocen que se debe realizar una redefinición operativa y conceptual de la población indígena, utilizando la pertenencia étnica. Sin embargo, comentan también que el enfoque de uso de la lengua porque los diversos recursos —culturales, económicos, sociales— que identifican al descendiente de un pueblo originario, se concentran alrededor del lenguaje (CONEVAL, 2014: 144).

<sup>12</sup> En México se ha cambiado en varias ocasiones el nombre de esta instancia. En 2002 cuando comienzan estos estudios se tiene al Instituto Nacional Indigenista (INI). Posteriormente en 2003, se pasa a la Comisión Nacional para el Desarrollo de los Pueblos Indígenas (CDI). Actualmente (2020), se denomina Instituto Nacional de los Pueblos Indígenas (INPI).

<sup>13</sup> El INPI (Instituto Nacional de los Pueblos Indígenas), unidad del gobierno especializada en la materia clasifica a los municipios con presencia indígena. Establece categorías utilizando el criterio de la lengua. Esta clasificación se encuentra vigente hasta la fecha (2020) y es referida en las diferentes reglas de operación y lineamientos relacionados con programadas dirigidos a pueblos originarios. Para más información véase: <http://www.inpi.gob.mx/transparencia/gobmxinpi/indicadores/>

Tras el análisis de las metodologías alternativas, se observa se generan importantes discrepancias en la cantidad de personas que pueden ser indígenas en México de acuerdo con la inclusión de la variable de autoadscripción (Rubio, 2014: 38). Por ejemplo, en 2010 esta población se estima en un rango que comprende desde poco más de 11 millones de personas (11'132,562), hasta poco más de 16 millones (16'102,646). La estimación de 11'132,562, corresponde al uso estricto de la lengua (INPI), mientras que el dato de 16'102,646 corresponde a la inclusión de datos de autoadscripción (INEGI). Si el último dato fuera el criterio de identificación, la población indígena podría haber sido 44.64 por ciento mayor a la que se definió en 2000.

La defensa del enfoque de identificación utilizando la lengua se justifica en la facilidad numérica de la variable (Valdés, 1995: 17). Según Navarrete (2010: 151), permite separar de forma sencilla a las personas que hablan castellano (y por tanto son mestizas) de las personas que conservan las lenguas originarias (y deberán ser indígenas). A pesar de esta situación, se ha debido incorporar el indicador cultural de autoadscripción, sin que ello implique una vinculación de facto con la política pública. Todo parece indicar que se incluye la autoadscripción como una estrategia para cumplir con el Convenio 169 de la OIT. La buena noticia es que la inclusión de esta variable permite realizar estudios estadísticos alternativos de la población indígena. En la siguiente sección, se discuten la forma en que, en la literatura, se analiza la relación autoadscripción y uso de la lengua.

### **USO DE LA LENGUA Y AUTOADSCRIPCIÓN EN LOS DATOS OFICIALES**

Existen una serie de trabajos que ya han utilizado esta información principalmente para explicar el comportamiento de las variables uso de la lengua y autoadscripción (Partida, 2005; Corona y Tuirán, 2001; Janssen y Martínez, 2006; Vázquez-Sandrín y Quesada, 2015).

En Partida (2005) se han realizado estimaciones de la población indígena, tomando al hogar y la relación de parentesco como fuentes importantes de variación en la cantidad de personas indígenas en México. No sólo se incluye a los hablantes de lengua, o a quienes se adscribieron como pertenecientes a un grupo autóctono en los censos, sino también a todos los miembros del hogar, cuando alguno de sus integrantes habla alguna lengua indígena, o se considera perteneciente a alguna etnia nativa. El estudio demuestra que de acuerdo con esta definición podrían existir más personas indígenas que las se reportan con el criterio de uso de la lengua. Esta cantidad es cercana a los 11'897,010

En Janssen y Martínez (2006) se tiene una definición similar de indígena, a la de Partida. Se incorpora la dimensión familiar y del hogar a las estimaciones extendiendo a todos los integrantes la característica indígena si por lo menos uno de ellos se define como tal y este integrante en el hogar no es un empleado doméstico (denominada definición extensa). Estos mismos autores plantean una tercera posibilidad que extiende la característica de indígena sí y solo sí, por lo menos una persona mayor de 15 años se define como indígena, o la jefatura del hogar o su cónyuge se definen como indígenas (denominada definición restringida). Los cálculos, muestran que para el año 2000 muestran la posible existencia de 12.7 millones de indígenas. Si se utiliza la definición extensa, la estimación de Janssen y Martínez tiene un valor exacto de 12'687,307 y será de 11'433,196 si se utiliza la definición restringida (por debajo de la estimación de Partida).

El análisis del siguiente censo en México (2010) también es interesante, debido a que existió un importante crecimiento en la población autoadscripta como indígena. De 2000 a 2010, la cifra casi se triplicó y representa la mayor proporción de población indígena desde la primera vez que se usó una variable de autoadscripción en 1932. Esta situación pudo ocurrir por diferentes razones, pero existe la sospecha de que la inclusión de la pregunta de autoadscripción fue determinante.

Vásquez Sandrin y Quesada (2015) señalan que Patricia Fernández en 2011 indica que en México se observó una “revitalización de lo étnico” de 2000 a 2010. Este fenómeno consiste en que hay un aumento de las personas que declararon ser indígenas, pero no hablan una lengua indígena, tal cual ya lo había planteado Barabas (2004: 2). Se mencionan como posibles causas de la revalorización de lo étnico dos situaciones: i) el reconocimiento de derechos de los pueblos indígenas (*e.g.* firma de C169) y; ii) la aparición del movimiento zapatista.

Sin embargo, en Vázquez-Sandrin y Quesada (2015), también se tiene una interpretación diferente del aumento de la población que se autoadscribe: el crecimiento en la cantidad de personas que se autoidentifican como indígenas, en realidad se debe al fraseo de la pregunta realizada en 2010, el cual “fue más permisivo”.

La hipótesis se relaciona con las ideas de Peyser y Chackiel (1999) en donde es posible una sobredeclaración en la autoadscripción por una “adhesión de simpatizantes a la causa indígena” de personas que no necesariamente son indígenas. Al mismo tiempo, ciertas preguntas como “pertenencia a pueblos indígenas y a cuáles de ellos” podrían generar una

subdeclaración al considerarse ofensivas o discriminatorias para quien responde.

Para ello, se evaluó la calidad de los instrumentos censales, aplicando los índices de Whipple, de Myers (United Nations, 1955: 40). Posteriormente prueban otras fuentes de variación en la respuesta, tales como los cambios en las tasas de fecundidad, la escolaridad y la migración. Mediante la eliminación progresiva de cada una de estas variables como hipótesis nulas, concluyen que la explicación más consistente es el efecto de fraseo en la pregunta (Vázquez-Sandrin y Quesada, 2015: 202). También concluyen que, si bien existen diferencias específicas en la calidad entre los censos de 2000 y 2010, estas diferencias no comprometen los instrumentos (Vázquez-Sandrin y Quesada, 2015: 186). Sin embargo, la metodología empleada en realidad no prueba estadísticamente la hipótesis del fraseo. Solo se mantiene como una posible hipótesis alternativa en torno a las variaciones entre uso de la lengua y autoadscripción, reconociendo que otra posible causa puede ser la revitalización de lo étnico (Vázquez-Sandrin y Quesada, 2015: 210).

Las diferentes características socioeconómicas y del contexto del hogar, podrían estar relacionadas con las discrepancias entre la variable de autoadscripción y uso de la lengua. Sin embargo, la evidencia es débil a lo largo del tiempo, debido a que las diferencias estadísticas entre los grupos de análisis parecen reducirse. Tampoco ha sido posible conocer cuáles son las causas específicas por las que se generan estas discrepancias. No se ha encontrado evidencia de que los instrumentos sean de mala calidad. Sin embargo, se asume que el fraseo de la pregunta contribuye a la discrepancia de información. Al mismo tiempo, es factible un proceso de revitalización de lo étnico. En este caso no existen datos para realizar alguna prueba estadística. Por fortuna, a lo largo del tiempo también se han generado instrumentos de captación de información más detallados que permiten conocer a nivel micro dato las características de las personas para construir modelos que permitan realizar inferencias más robustas. A continuación, se presenta un ejercicio estadístico en el que se arroja luz sobre las causas de estas discrepancias estadísticas entre uso de la lengua y autoadscripción, con información de la Encuesta Intercensal 2015.

### **METODOLOGÍA PARA ANALIZAR LA RELACIÓN ENTRE USO DE LA LENGUA INDÍGENA Y AUTOIDENTIFICACIÓN**

Inicialmente se proponen dos modelos: uno para explicar la probabilidad de que una persona indique que habla la lengua indígena, y otro, donde la

variable dependiente será la probabilidad de que la persona se autoadscribe como indígena. En este caso, la técnica de estimación de los parámetros de cada modelo (*i.e.* los impactos de cada variable) será mediante un modelo donde se asume un proceso de autoselección entre las variables de uso de la lengua indígena y autoidentificación como indígena. La estimación de parámetros en este caso es mediante un modelo de corrección de sesgo de autoselección de Heckman (1979). En la sección “Formalización de los modelos” se indica con más detalle las razones técnicas de utilizar estos dos modelos.

### Sobre los datos que se utilizan

La Tabla 1, resume la historia de los principales instrumentos donde se puede conocer la evolución de la población indígena de 2000 a 2015.

Tabla 1: Evolución de la población indígena de 2000 a 2015 en México

	Habla LI	No habla LI	Total	Grupo poblacional
Se considera (2000)	4'151,753 4.96%	1'101,316 1.32%	5'258,852 6.28%	5 o más años
Se considera (2010)	6'553,954 6.20%	9'141,099 8.65%	15'701,279 14.86%	3 o más años
Se considera (2015)	6'711,553 5.90%	17'817,355 15.67%	24'556,327 21.60%	3 o más años

Fuente: elaboración propia con datos de INEGI para diferentes años.

En 2015 se realiza la Encuesta Intercensal,<sup>14</sup> la cual permite conocer que las personas que hablan una lengua indígena y se autoadscriben suman un total de 6'711,553 (alrededor del 5.90 por ciento).

Sin embargo, el grupo de personas que se autoidentifican como indígenas en este mismo año, se compone por 24'556,327 que equivale alrededor de 21.6 por ciento. Al comparar las dos cifras se observa que quienes se autoadscriben son seis veces más que los que hablan la lengua. Un total de 17'817,355 (alrededor de 15.67 por ciento) se considera indígena, aunque no hable alguna lengua indígena. Este conjunto de personas en México para 2010 se componía por 9'141,099 personas y representaba 8.65 por ciento. El mismo grupo en 2000 eran de apenas 1'101,316 que representaba 1.32 por ciento por ciento.

<sup>14</sup> Véase: [http://internet.contenidos.inegi.org.mx/contenidos/productos//prod\\_serv/contenidos/espanol/bvinegi/productos/nueva\\_estruc/702825078966.pdf](http://internet.contenidos.inegi.org.mx/contenidos/productos//prod_serv/contenidos/espanol/bvinegi/productos/nueva_estruc/702825078966.pdf)

¿Por qué este dato es de suma importancia? Si se asume que el criterio de identificación prevaleciente en 2015 es el uso de la lengua indígena, se pudieron haber omitido a 17'817,355 de personas, como descendientes de un pueblo indígena. Esta cantidad de personas equivale a la población de cualquiera de los estados del país en 2015, incluyendo al Estado de México que es la entidad con mayor población reportada (16'187,608 personas). Adicionalmente, implica que en realidad el estado mexicano tuvo que generar seis veces más acciones para la población indígena y destinar más presupuesto para ello del que destina conforme a la definición del uso de la lengua.

Dada la importancia del estudio de la discrepancia entre autoadscripción y uso de la lengua, en el siguiente trabajo se han realizado ejercicios estadísticos con los micro datos de la Encuesta Intercensal 2015. Inicialmente se realizaron los cálculos con una muestra aleatoria de 156,579 observaciones debido a que con ello se favorece la generación de cálculos de forma rápida. En particular para el modelo de corrección de autoselección. Posteriormente, se replicaron las estimaciones con la muestra completa (6'133,946),<sup>15</sup> sin que se mostraran diferencias considerables en signo y magnitud, para los parámetros estimados con la muestra.

### **Formalización de los modelos**

La decisión de autoadscribirse o bien indicar el uso de la lengua, es afectada por situaciones particulares de las personas que no se observan. Se sospecha de este sesgo debido a como observamos en los datos y análisis previo, personas que no hablan la lengua indígena se están identificando como personas indígenas de forma sistemática. Esta condición podría aludir a un efecto de sobrerepresentación en la muestra de una manera deliberada (Cameron y Trivedi, 2005).

Dadas las características del diseño metodológico del INEGI (Encuesta Intercensal 2015,<sup>16</sup> Censo de 2010, 2005 y la muestra del Censo de 2000), no es posible argumentar que existe un sesgo intencional de la sobrerepresentación por parte de Instituto. En los trabajos previamente analizados, se asume que el sesgo pudo ser ocasionado por un fenómeno denominado revitalización de lo étnico o por un error de fraseo.

<sup>15</sup> Véase: [http://internet.contenidos.inegi.org.mx/contenidos/Productos/prod\\_serv/ contenidos/espanol/bvinegi/productos/nueva\\_estruc/702825078836.pdf](http://internet.contenidos.inegi.org.mx/contenidos/Productos/prod_serv/ contenidos/espanol/bvinegi/productos/nueva_estruc/702825078836.pdf)

<sup>16</sup> Véase: [http://www.inegi.org.mx/est/contenidos/proyectos/encuestas/hogares/ especiales/ei2015/doc/eic2015\\_sintesis.pdf](http://www.inegi.org.mx/est/contenidos/proyectos/encuestas/hogares/ especiales/ei2015/doc/eic2015_sintesis.pdf)

Hasta donde se ha revisado, en la literatura no se ha planteado cómo es la causalidad entre indicar el uso de la lengua indígena y la autoadscripción (Gibson, 1997; Smith, 1997: 18; Moreno y Oropesa, 2012; Moreno, 2014).

En el trabajo de Romer (2010), se encuentra un argumento que podría arrojar luz en torno a la relación de estas variables: la pérdida de la lengua, no implica necesariamente la pérdida de identidad. Sostiene que los padres no enseñan la lengua a sus hijos desde pequeños, porque tienen como objetivo el que los niños indígenas, al contar un adecuado conocimiento del español, “superen su condición de indígena” al acceder a niveles más altos de escolaridad (Romer, 2010: 214).

Por otro lado, se puede asumir que las personas pueden identificarse como indígenas, debido a que al hacerlo facilita el ser candidato a ciertos programas sociales (Perraudin, 2010). En ambos casos el sesgo proviene de una decisión de las personas.

En ausencia de pruebas empíricas previas, se ha optado entonces por modelar el uso de lengua con una ecuación de autoselección donde aparece la autoadscripción, pero también, la autoadscripción con una ecuación de autoselección donde aparece el uso de la lengua.

Si existe un sesgo de selección, es necesario corregirlo y para ello se utilizar un modelo de Heckman (1979). En particular, se utilizará un Modelo Probit<sup>17</sup> con corrección de sesgo de selección (Wooldridge, 2002: 570-571). El método consiste en estimar en un primer paso, la probabilidad de que una persona decida o no indicar que habla una lengua indígena.

Posteriormente, se obtiene un estadístico denominado razón inversa de Mills, que es la razón de la función de densidad entre la función de distribución acumulada. Esta razón es importante, porque captura la magnitud del sesgo (Tobin, 1958; Heckman, 1976; 1979). Derivado del cálculo del primer modelo Probit, la razón inversa de Mills se incluye en una segunda regresión como un regresor más. Si el coeficiente resulta ser estadísticamente diferente de cero, entonces se puede interpretar su magnitud y signo como el sesgo en que se incurre.

Generalmente la segunda regresión en un modelo de Heckman se hace mediante una estimación de Mínimos Cuadrados Ordinarios (MCO) y de forma conjunta (*i.e.* ambas regresiones al mismo tiempo). Sin embargo, dado que la variable a explicar en esta segunda regresión también es de respuesta binaria, nuevamente aplicaremos un modelo Probit para la estimación de los parámetros del vector de variables independientes que corrigen el sesgo en esta ecuación.

<sup>17</sup> Para los detalles sobre este tipo de modelo ver Wooldridge (2002).

Formalmente:

$$x_1^* = x_1\beta_1 + \varepsilon_1$$

$$y = x_2\beta_2 + \varepsilon_2 \text{ si } x_1^* > 0$$

$$y \text{ no se observa si } x_1^* \leq 0$$

La primera ecuación es la estimación sobre una variable latente  $x_1^*$  que no se observa. Las variables  $x_1$  y  $x_2$  son vectores de características de las personas, mientras que  $\beta_1$ ,  $\beta_2$  serán vectores de parámetros. Es importante notar que en este modelo la variable  $y$ , se encuentra en función de un conjunto de variables  $x_2^*$  pero se observa si y solo si, al mismo tiempo la variable latente  $x_1^*$  (*i.e.* la variable de selección) también se observa. En otras palabras: la variable dependiente  $y$ , es función del valor de otra regresión sobre la variable  $x_1^*$ , por lo cual, también se relaciona con algunas características  $x_1$ .

Si se asume que los errores siguen una distribución normal bivariada con media 0 y varianza  $\sigma_1^2$ , tendremos que:

$$\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix} \sim N \left[ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho\sigma_2^2 \\ \rho\sigma_2^2 & \sigma_2^2 \end{pmatrix} \right]$$

Lo anterior es posible al aplicar el supuesto de normalidad, normalizar la varianza de la ecuación de selección a uno y aplicar las propiedades de la función de densidad normal bivariada trunca. En esta matriz varianza covarianza de los errores del sistema de ecuaciones, existe una correlación  $\rho$  entre  $\varepsilon_1$  y  $\varepsilon_2$ . Esta correlación será importante debido a que es la que permite conocer el sesgo de selección, en particular el signo.

Asumiendo los supuestos anteriores, obtenemos una ecuación de selección, que en la práctica se convierte en un modelo Probit y que permite definir el valor esperado de la variable dependiente condicionado en los valores de  $x_1^*$  como:

$$E(y|x_1^* > 0) = x_2\beta_2 + E(\varepsilon_2|\varepsilon_1 > -x_1\beta_1)$$

Tendremos que  $E(y|x_1^* > 0)$  puede ser expresado en función de un vector de variables  $x_2$  así como de un término que incluye el sesgo ocasionado por la ecuación de selección ( $E(\varepsilon_2|\varepsilon_1 > -x_1\beta_1)$ ).<sup>18</sup>

<sup>18</sup> Para revisar el procedimiento a detalle se recomienda el texto de Wooldridge (2002: 560-571).

$$E(y|x_1^* > 0) = x_2\beta_2 + E(\varepsilon_2|\varepsilon_1 > -x_1\beta_1)$$

Donde  $P(y|x_1^* = 1) = \phi(x_1\beta_1)$  es el modelo probit que permite estimar  $\hat{\beta}_1$ . Si la razón inversa de Mills y la varianza  $\sigma_2^2$  siempre son positivas, el signo del sesgo de selección que aparece en el segundo término del segundo miembro de la ecuación se determina por el valor de  $\rho$ . Al incluir el sesgo de selección, la magnitud del impacto del sesgo de selección depende de la correlación entre los errores de las ecuaciones iniciales, la varianza de la ecuación se que modela, la variable a explicar y la forma de la estimación del inverso de Mills, en particular  $\phi(x_1\beta_1)$ . Si  $x_1\beta_1$  es muy chico, es claro que el sesgo de selección independientemente del valor de  $\rho$  será menor.

Para los ejercicios estadísticos se tiene entonces dos grupos de modelos de la siguiente forma:

1) Uso de la lengua en la ecuación de interés y autoadscripción en la ecuación de selección.

$$\text{Autoadscripción}^* = x_1\beta_1 + \varepsilon_1$$

$$\text{Uso de la lengua} = x_2\beta_2 + \varepsilon_2 \text{ si } x_1^* > 0$$

$$\text{Uso de la lengua no se observa si } \text{Auto}^* \leq 0$$

2) Autoadscripción en la ecuación interés y uso de la lengua en la ecuación de selección.

$$\text{Uso de la lengua}^* = x_1\beta_1 + \varepsilon_1$$

$$\text{Autoadscripción} = x_2\beta_2 + \varepsilon_2 \text{ si } x_1^* > 0$$

$$\text{Autoadscripción no se observa si } \text{Uso de la lengua}^* \leq 0$$

A continuación, se comentan las hipótesis y variables utilizadas.

### Hipótesis de interés

En los estudios previamente citados, se utiliza principalmente el contexto del hogar para explicar el uso de la lengua indígena y la autoadscripción. Por lo tanto, la introducción de variables que caractericen el hogar de una persona es imprescindible para entender la relación entre autoadscripción y uso de la lengua.

En relación con la primera hipótesis, se plantea que la escolaridad de la jefatura del hogar incide en mayor medida que el uso de la lengua y en la

autoadscripción. En este caso se espera que la beta asociada sea estadísticamente diferente de cero y con signo negativo.

Para la segunda hipótesis, se busca probar que en los hogares que reciben apoyos por parte del gobierno, se incentiva a las personas a declararse indígenas o indicar el uso de la lengua. Por lo tanto, se espera que la beta asociada sea estadísticamente diferente de cero y con signo positivo.

## **VARIABLES**

### **Variables dependientes**

Se plantean dos modelos. En el primero, la variable dependiente es el uso de la lengua. En el segundo modelo, se presenta ahora como variable dependiente la autoidentificación. Ambas variables identifican con un valor de 1 la existencia del atributo en cada caso y 0 en caso de que no exista el atributo en cuestión. Más adelante se detalla cómo se integran estas variables para corregir el sesgo de selección.

### **Variables independientes**

Para probar la primera hipótesis se tiene la escolaridad de la jefatura de familia, pero también la educación de la persona (ambas variables discretas, en años). Para la segunda hipótesis, se plantea que los ingresos derivados de apoyos del gobierno inciden positivamente en declarar el uso o en autoadscribirse. En este caso se tiene una variable de respuesta binaria que indica con el valor de 1 si en el hogar existe el ingreso en cuestión.

Si la decisión de identificarse como indígena depende de los apoyos del gobierno también es posible que dependa de los ingresos que aportan las personas que pertenecen al hogar, pero que están en el extranjero.

Existen dos razones para sospechar que esto ocurre: i) en México, una cantidad importante de personas que tienen familiares en el extranjero, provienen de lugares con una fuerte presencia indígena; ii) el obtener recursos del extranjero, pudiera rivalizar con el recibir ingresos por parte del gobierno. Esto implicaría que en la medida en que los hogares se vuelven menos dependientes del gobierno, encuentran conveniente dejar “su condición de indígenas”. Para separar estos efectos, se incluye una variable de respuesta binaria que identifica con el valor de 1 si el hogar recibe ingresos de personas que viven en otro país.

## Variables de control

Se tienen varios grupos de variables de control. En el primer grupo están características de la persona tales como: i) la persona habla español (respuesta binaria); ii) escolaridad (discreta, en años); iii) la persona es mujer (respuesta binaria) y; iv) edad de la persona (discreta, en años).

Un segundo grupo de variables se relacionan con el contexto del hogar. Se plantean las siguientes variables de respuesta binaria en caso de que se tenga el atributo: i) el hogar es administrado por una mujer; ii) la jefatura de hogar se autoadscribe como indígena; iii) en el hogar quien tiene la jefatura habla español; iv) la jefatura del hogar habla una lengua indígena; v) edad de la jefatura de familia (discreta, en años) y; vi) número total de integrantes del hogar (discreta).

Finalmente, se incluyen 31 variables de respuesta binaria en las que se captura el efecto que pueden tener aspectos particulares asociados a las entidades de México. Se establece como grupo de control la Ciudad de México. No se reportan los parámetros de estas variables por cuestiones de espacio.

## RESULTADOS

En la Tabla 2 aparecen las estimaciones de los coeficientes para el modelo en la ecuación de selección (*i.e.* Autoadscripción = Auto\*, Uso de la lengua = Uso\*) que es la que permite identificar la existencia de un sesgo. Se utiliza el método Delta<sup>19</sup> para estimar la varianza, errores estándar e intervalos de los parámetros.

Para el modelo de uso de la lengua (donde la ecuación de autoselección es la autoadscripción), el parámetro de la variable donde la jefatura se autoadscribe es el más alto (0.295). El mismo parámetro es mucho más pequeño (0.057) en el modelo de la autoadscripción (con ecuación de autoselección uso de la lengua). Por otro lado, si la jefatura del hogar habla una lengua indígena, en el modelo de uso de la lengua (*i.e.* con la variable de autoadscripción como dependiente) el impacto es de 0.066 mientras que en el modelo de la autoadscripción (con uso de la lengua como dependiente en la ecuación de autoselección) es de 0.184.

Con estos valores estimados se puede inferir que para una persona que habla la lengua, resulta fundamental que la jefatura de su hogar se autoadscribe como indígena. Al mismo tiempo, si una persona se autoadscribe es

<sup>19</sup> Long (1997) indica que la estimación de los efectos marginales en las medias tiene limitaciones por lo que es más recomendable, especialmente en el caso de variables binarias calcular el cambio discreto de estas.

importante que la jefatura de hogar indique el uso de la lengua indígena. Sin embargo, el efecto de la autoadscripción de la jefatura del hogar sobre el uso de la lengua es más poderoso, que el efecto del uso de la lengua de la jefatura del hogar, sobre la autoadscripción de las personas.

Tabla 2: Resultados de la ecuación de selección

	Lengua (Autosel.= Autoadscripción)	Autoadscripción (Autosel. = Lengua)
<i>Contexto del hogar</i>		
Una mujer tiene la jefatura del hogar (1,0)	-0.010 ***	-0.012 ***
Edad de la jefatura del hogar (1,0)	-0.001 ***	-0.001 ***
Escolaridad de la jefatura del hogar (0, ...24)	-0.001 ***	-0.003 ***
La jefatura del hogar se autoadscribe (1,0)	0.295 ***	0.057 ***
La jefatura del hogar habla una lengua (1,0)	0.066 ***	0.184 ***
Número de integrantes (0, ...,38)	0.000	0.002 ***
<i>Características de la persona</i>		
La persona es mujer (1,0)	0.003 **	0.006 ***
Edad de la persona (0, ...24)	0.001 ***	0.002 ***
Escolaridad de la persona (1,0)	-0.001 ***	-0.001 ***
<i>Comportamiento estratégico</i>		
El hogar recibe ingresos del gobierno (1,0)	0.018 ***	0.018 ***
El hogar recibe ingresos de migrantes (1,0)	-0.005 **	-0.008 ***

Nota 1: se utiliza para la determinación de p-values lo siguiente \* para 10 por ciento, \*\* para 5 por ciento y \*\*\* para significancia de uno por ciento y menos.

Nota 2: entre paréntesis se incluyen los rangos de valores de cada variable. El tipo de variable que más se repite es la de respuesta binaria identificada como (1,0).

Fuente: elaboración propia con datos de la Encuesta Intercensal 2015 de INEGI.

Si la persona es mujer y jefe en el hogar, hay una menor probabilidad tanto de indicar el uso de la lengua, así como de autoadscribirse. Mayor edad de la jefatura del hogar reduce la probabilidad de que una persona indique el uso de la lengua indígena, así como de que se autoadscribe como indígena.

En relación con la escolaridad de la jefatura de familia, un año adicional, tiene un impacto negativo en que las personas se autoadscriben como indígenas y hablen la lengua indígena. Mayor número de integrantes impacta de manera positiva la probabilidad de que la persona se autoidentifique como indígena.

Para las variables relacionadas con atributos de la persona, se tiene un aumento en la probabilidad para el uso de la lengua y la autoadscripción, cuando el entrevistado es mujer y cuando tiene una mayor edad. Estos impactos, contrastan con los impactos de las variables que identifican lo mismo, pero en el caso de la jefatura de hogar donde se observó exactamente el signo contrario. Este resultado es interesante y sugiere una posible disyuntiva en la identidad entre las personas que viven en un hogar y las jefaturas de este. Dicho de otra forma: en los hogares que se asumen como indígenas no todos sus integrantes se consideran como tal por lo cual no deberían considerarse como unidades homogéneas.

Se encuentra que una mayor escolaridad de la persona impacta negativamente en ambos modelos. Esta estimación está en consonancia con los resultados previos para el caso de la jefatura de hogar y permite inferir que a medida que se tengan hogares con una mayor escolaridad en el país, se generarán incentivos para que las personas decidan no identificarse como indígena y declarar el uso de la lengua. Por lo anterior, se confirma que efectivamente una mayor escolaridad en el hogar impacta negativamente en la intención de autoadscribirse e indicar el uso de la lengua indígena.

En cuanto a la hipótesis del comportamiento derivado de los ingresos, se observa que, si el hogar recibe ingresos de programas de gobierno, aumenta la probabilidad de que la persona se autoadscribe y de que declare hablar una lengua indígena. Al mismo tiempo, se identifica que si el hogar recibe ingresos de personas que viven en otro país, aparece una reducción en la probabilidad de que la persona use la lengua indígena, así como de que se autoadscribe como indígena. En otras palabras: depender del gobierno incentiva a indicar hablar la lengua indígena y autoadscribirse, mientras que recibir ingresos de familiares en el extranjero reduce la intención de que la persona se identifique o indique ser indígena.

A manera de conclusión de los resultados de los modelos, se puede decir que con los datos oficiales se revela un proceso de lucha interna por la identidad en los hogares en el cual el acceso a mejores condiciones de ingreso y educación resulta crucial. La condición de educación y el recibir ingresos de familiares en el extranjero, incentiva a las personas al abandono de la “condición de indígena”. La existencia de apoyos del gobierno propicia que las personas quieran identificarse como indígenas ya que encuentran útil tal situación.

### **INDEPENDENCIA ENTRE LA AUTOADSCRIPCIÓN Y EL USO DE LA LENGUA INDÍGENA**

Con los datos del modelo de autoselección, se puede estimar un estadístico que permite probar la independencia de ecuaciones (*i.e.* ecuación de interés y de selección). Se podría pensar que no es posible separar el uso de la lengua de la autoadscripción pues son variables equivalentes. Hasta donde se conoce, esta hipótesis no se ha probado estadísticamente para el caso de México. Simplemente, se asume una cierta relación entre las variables. En la Tabla 3 se presentan información de una prueba LR diseñada específicamente para probar independencia entre la ecuación de interés y la de selección.

El estadístico que se calcula, se denomina LR (Likelihood-Ratio, razón de verosimilitud) y establece como hipótesis nula, la independencia en las ecuaciones de interés y de autoselección. Si el p-value asociado al estadístico LR es muy pequeño (cercano a 0), entonces diremos que las ecuaciones no son independientes y es necesario modelar la autoselección. El valor de  $r$  que indica el sesgo de selección.

De la información que se presenta, se observa un estadístico LR ( $r = 0$ ):  $\chi^2$  de 15.350 (p-value de 0.000) para el modelo que explica a la lengua indígena y de 1.890 (p-value de 0.170) para el modelo donde la variable dependiente es la autoadscripción. Por lo tanto, se rechaza la hipótesis de que el uso de la lengua es independiente de la autoadscripción, pero no se rechaza que la autoadscripción sea independiente del uso de la lengua.

Se infiere que la autoidentificación es un proceso independiente del declarar el uso de la lengua; sin embargo, el que una persona indique el uso de la lengua, no es independiente de que una persona se autoadscribe como indígena. El coeficiente de  $\rho$  resulta ser (como se esperaba) positivo y estadísticamente diferente de cero solamente para el caso del modelo de uso de la lengua (p-value 0.001). Este dato confirma que efectivamente hay un

sesgo de autoselección positivo, provocado por la autoadscripción al explicar la variable uso de la lengua. De forma más sencilla: en la medida que la persona decide autoadscribirse es más probable que indique el uso de la lengua. El recíproco no es verdadero.

Tabla 3: Pruebas de independencia

	Lengua	Autoadscripción
Observaciones	156,579	156,579
Observaciones censuradas	99,689	132,277
Observaciones no censuradas	56,890	24,302
Wald $\chi^2(40)$	21,785.640	2167.770
Prob > $\chi^2$	0.000	0.000
Log likelihood	-47837.750	-22601.800
$\rho$	0.279	0.155
Prob > $ z $ de $\rho$	0.001	0.195
Prueba LR de ecuaciones independientes ( $r = 0$ ): $\chi^2$	15.350	1.890
Prob > $\chi^2$	0.000	0.170

Nota: los grados de libertad en el estadístico Wald presentan una cantidad mayor de variables debido a que se incluyen otras variables no reportadas en los resultados finales. Estas variables como se ha comentado son 31 y corresponden a los estados de la República Mexicana.

Fuente: elaboración propia con datos de la Encuesta Intercensal 2015 de INEGI.

## CONCLUSIONES Y DISCUSIÓN

En el presente trabajo se ha realizado un esfuerzo por demostrar que las variables de uso de la lengua y autoadscripción están relacionadas, pero no son iguales. Se concluye que la variable de uso de la lengua indígena es un subconjunto de la variable de autoadscripción. Esto implica que ambas variables no pueden ser considerados como sinónimos (o conceptos equivalentes).

Los resultados del modelo de autoselección demuestran que cuando se explica el uso de la lengua en función de la autoadscripción, se rechaza su independencia. Por otro lado, cuando se explica la autoadscripción en función de la lengua no se rechaza la independencia. En síntesis: el uso de la lengua está condicionado por la autoadscripción, pero la autoadscripción no se condiciona por el uso de la lengua. Por lo anterior, la variable

relevante para identificar a la población indígena, no debería ser el uso de la lengua, sino la autoadscripción. Las personas que se autoadscriben, incluirán al subconjunto de quienes indican usar la lengua (o al menos a la mayoría de dicho subconjunto tal cual se ilustra en el esquema). Por otro lado, el autoadscribirse no tiene por qué estar vinculado al uso de la lengua, si bien podría existir una intersección entre las dos variables.

Una aportación de este trabajo es explicar por qué existe dicho sesgo. Los resultados de las estimaciones de los modelos estadísticos soportan la evidencia de que, si el hogar recibe ingresos de programas de gobierno, la persona tendrá una motivación de declarar el uso de la lengua o autoadscribirse. Contar con programas especializados destinados a las personas consideradas como indígenas, puede motivar un incremento en la población indígena. Esto no implica que puedan existir personas que se autoadscriben, sin que se racionalice la obtención de un recurso público. Sin embargo, las estimaciones confirman que el sesgo en las estadísticas proviene de un comportamiento estratégico.

Por otro lado, se ha encontrado que cuando los hogares tienen fuentes de ingresos diferentes a las del gobierno, como los ingresos de personas que viven en el extranjero, la probabilidad de que en ese hogar las persona indiquen el uso de la lengua y la autoadscripción se reduce. De este resultado se concluye que mejorar el ingreso por una vía que no es la transferencia del gobierno, hace que las personas decidan abandonar su identidad indígena.

Se identificó que existe una fuerte influencia cuando la jefatura de hogar donde vive la persona declara hablar la lengua o se autoadscribe. Esto permite asumir una transferencia de valores culturales y de la identidad indígena de jefes de hogar a los demás integrantes. Sin embargo, también se observó un efecto en contra de dicha identidad cuando la persona es mujer y jefe en el hogar, a mayor edad de la jefatura del hogar y a mayor escolaridad tal cual. La influencia negativa de la escolaridad en las variables de autoadscripción y uso de la lengua, refuerza los hallazgos de Romer (2010).

Tras los hallazgos, se proponen algunas recomendaciones: i) incluir siempre que sea posible la autoadscripción en los instrumentos de captación del INEGI a nivel persona (encuestas, censos, conteos, módulos); ii) explorar la posibilidad de cambiar de manera aleatoria, el orden de las preguntas en los cuestionarios que aplica el INEGI para evitar que siempre exista un condicionamiento de la lengua hacia la autoadscripción en las personas entrevistadas (e.g. invertir el orden de las preguntas); iii) es imprescindible discutir en los libros y programas de primaria, las implica-

ciones específicas de la importancia de mantener la identidad a fin de evitar la motivación de “abandonar la condición de indígena” en las personas que adquieren mayores niveles de escolaridad: iv) desarrollar estudios específicos sobre los posibles sesgos en los programas sociales focalizados a la población indígena.

Finalmente, se plantea una última pregunta en materia de política pública: ¿debería la autoridad incluir oficialmente a las personas que se autoadscriben como indígenas como población indígena o debería mantener el criterio del uso de la lengua? Esta es una de las interrogantes que todavía abiertas y en espera de una pronta respuesta. Nuestra respuesta es que se deben incluir pese a la existencia de los sesgos que se han encontrado y dado el compromiso contraído por el estado mexicano al ratificar el C169. En todo caso, una vez que se conocer de estos sesgos, se debería tratar de incluirlos y no simplemente para evitarlos excluir a una parte importante de la población indígena del reconocimiento de su identidad.

## REFERENCIAS BIBLIOGRÁFICAS

- Barabas, A., 2004, “Un acercamiento a las identidades de los pueblos indios de Oaxaca”, en *Amérique Latine Histoire et Mémoire. Les Cahiers ALHIM*, 10, consultado el 24/04/2020. Disponible en <http://alhim.revues.org/105>
- Carrasco, T. y Alcázar, Tatiana, 2009, “Los pueblos indígenas y los censos en México y América Latina: La cultura en la definición de su identidad”, en Valdés, Luz María (coord.), *Derecho de los mexicanos; Introducción al derecho demográfico*, acervo Biblioteca Jurídica Virtual del Instituto de Investigaciones Jurídicas, UNAM, consultado el 24/04/2020. Disponible en <https://archivos.juridicas.unam.mx/www/bjv/libros/6/2638/18.pdf>
- Cifuentes, B., y Moctezuma, J.L., 2006, “The Mexican indigenous languages and the national censuses: 1970-2000”, in *Contributions to the Sociology of Language*, 91, 191.
- CONAPRED, 2010, *Encuesta Nacional sobre Discriminación en México ENADIS, Resultados generales*, México: Consejo Nacional para la Prevención de la Discriminación (CONAPRED), consultada el 17/04/2020. Disponible en <http://www.conapred.org.mx/redes/userfiles/files/Enadis -2010-RG-Accss-001.pdf>
- CONAPRED, 2011, *Documento informativo sobre la discriminación racial en México*, CONAPRED, México, consultado el 24/04/2020. Disponible en [https://www.conapred.org.mx/documentos\\_cedoc/Dossier%20DISC-RACIAL.pdf](https://www.conapred.org.mx/documentos_cedoc/Dossier%20DISC-RACIAL.pdf)
- CONAPRED, 2012, *Encuesta Nacional Sobre Discriminación en México 2010, Resultados Sobre Diversidad Cultural*, consultado el 17/04/2020. Disponible en <http://www.conapred.org.mx/userfiles/files/Enadis-DC-INACCSS.pdf>
- CONEVAL, 2014, *La pobreza en la población indígena de México, 2012*, Primera edición, julio de 2014, Consejo para la Evaluación de la Política en Desarrollo

Social (CONEVAL) consultado el 20/04/2020. Disponible en [http://www.coneval.org.mx/Informes/Coordinacion/INFORMES\\_Y\\_PUBLICACIONES\\_PDF/POBREZA\\_POBLACION\\_INDIGENA\\_2012.pdf](http://www.coneval.org.mx/Informes/Coordinacion/INFORMES_Y_PUBLICACIONES_PDF/POBREZA_POBLACION_INDIGENA_2012.pdf)

Corona, R., y Tuirán, R., 2001, La migración internacional desde y hacia México, en *La Población en México. Tendencia y Perspectivas Socio Demográficas hacia el Siglo XXI*, CONAPO/FCE, México.

Gamio, M., 1942, “Las características culturales y los censos indígenas”, en *América Indígena*, 2(3), 15-19.

Gibson, C., 1998, “We sing our home, we dance our land: indigenous self-determination and contemporary geopolitics in Australian popular music”, in *Environment and Planning D: Society and Space*, 16(2), 163-184.

González-Ulloa, P.A., 2013, “Recognition and cultural diversity: the new realities in Mexico”, in *Asian Journal of Latin American Studies* (2013) vol. 26 nº. 2: 67-88

Goodenough, W., 1971, *Cultura, lenguaje y sociedad. En Kahn, El Concepto de cultura*. Textos fundamentales. Barcelona: Anagrama.

Heckman, J.J., 1976, “The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models”, in *Annals of Economic and Social Measurement*. 5 (4): 475-492.

Heckman, J.J., 1979, “Sample selection bias as a specification error”, in *Econometrica. Journal of the Econometric Society* (47): pp.153-161.

INEGI, 2000, XII *Censo General de Población y Vivienda 2000. Base de Datos de la Muestra Censal*. Descripción de archivos, Instituto Nacional de Estadística y Geografía (INEGI), consultado el 21/04/2020. Disponible en [http://www.beta.inegi.org.mx/contenidos/proyectos/ccpv/2000/doc/fd\\_muestra\\_censal\\_cgpv2000.pdf](http://www.beta.inegi.org.mx/contenidos/proyectos/ccpv/2000/doc/fd_muestra_censal_cgpv2000.pdf).

INEGI, 2004, *La población indígena en México. INEGI, México 2004*, consultado el 21/04/2020. Disponible en [http://internet.contenidos.inegi.org.mx/contenidos/productos/prod\\_serv/contenidos/espanol/bvinegi/productos/historicos/2104/702825497583/702825497583\\_1.pdf](http://internet.contenidos.inegi.org.mx/contenidos/productos/prod_serv/contenidos/espanol/bvinegi/productos/historicos/2104/702825497583/702825497583_1.pdf)

Janssen, E., y Martínez-Casas, R., 2006, “Una propuesta para estimar la población indígena en México a partir de los datos censales”, en *Estudios Demográficos y Urbanos*, 457-471.

Knight, A., 2010, “Racism, revolution and indigenismo: México 1910-1940”, en Graham, Richard (ed.), *The idea of Race in Latin America: 1870-1940*, University of Texas Press, Austin, págs. 71-113.

Landázuri, G., 2012, “La migración, marco de exclusión y discriminación en San Gregorio Atlapulco, México”, en Alicia Castellanos y Gisela Landázuri (coords.), *Racismos y otras formas de intolerancia de Norte a Sur en América Latina*, UAM, México, págs. 77-98.

Long, J.S., 1997, *Regression models for categorical and limited dependent variables*. Thousand Oaks, CA: Sage Publications.

- Molina, V., 2010, “Inserción laboral de los indígenas en la ciudad de México”, en Durin, Séverine (coord.), *Etnicidades urbanas en las américas. Procesos de inserción, discriminación y políticas multiculturalistas*, México, Publicaciones de la Casa Chata, págs. 77-94.
- Moreno, M., 2014, “Patrones de autoidentificación etnoracial de la población indígena en las encuestas de hogares en el Perú”, en *Debates en Sociología*, 39.
- Moreno, M. y Oropesa, R. S., 2012, “Ethno-racial identification in urban Peru”, in *Ethnic and Racial Studies*, 35(7), pp. 1220-1247.
- OIT, 2009, *Los Derechos de los Pueblos Indígenas y Tribales en la Práctica - Una Guía sobre el Convenio No. 169 de la OIT*. Programa para promover el Convenio núm. 169 de la OIT (PRO 169), Departamento de Normas Internacionales del Trabajo, Organización Internacional del Trabajo (OIT).
- ONU, 1955, *Manuals on methods of estimating population MANUAL II Methods of Appraisal of Quality of Basic Data for Population Estimates*, United Nations Publications, Organización de las Naciones Unidas (ONU), Nueva York.
- Partida, V., 2005, *Proyecciones de indígenas de México y de las entidades federativas 2000-2010*, consultado el 18/04/2020. Disponible en [http://www.conapo.gob.mx/work/models/CONAPO/indigenas\\_2010/Proyindigenas.pdf](http://www.conapo.gob.mx/work/models/CONAPO/indigenas_2010/Proyindigenas.pdf)
- Perraudin, A., 2010, “Estrategias residenciales e interacciones interétnicas en los asentamientos colectivos. El caso de los otomíes de Santiago Mexquititlán en la Ciudad de México”, en Durin, Séverine (coord.), *Etnicidades urbanas en las américas. Procesos de inserción, discriminación y políticas multiculturalistas*, México, Publicaciones de la Casa Chata, págs. 155-178.
- Pla, D., 2011, “Más desindianización que mestizaje. Una relectura de los censos generales de población”, en *Dimensión Antropológica*, 18, págs. 69-92.
- Romer, M., 2010, “Socialización, identidad y estigma. El caso de los hijos de inmigrantes indígenas en la Ciudad de México”, en Durin, Séverine (coord.), *Etnicidades urbanas en las américas. Procesos de inserción, discriminación y políticas multiculturalistas*, México, Publicaciones de la Casa Chata, págs. 207-226.
- Saldívar, E. y Walsh, C., 2014, “Racial and ethnic identities in Mexican statistics”, in *Journal of Iberian and Latin American Research*, 20(3), 455-475.
- Schkolnik, S., y Del Popolo, F., 2005, “Los censos y los pueblos indígenas en América Latina: una metodología regional”, en *Notas de Población*, 31(79), 101-32.
- Schmelkes, S., 2013, “Educación y pueblos indígenas: problemas de medición”, en *Revista Internacional de Estadística y Geografía*, 1, 5-13.
- Smith, A., 1997, *La identidad nacional*, Madrid, Trama Editorial.
- Stavenhagen, R., 1992, “La situación y los derechos de los pueblos indígenas de América, en *América Indígena*, núm. 1-2, México, Instituto Indigenista Interamericano, pp. 63-118.
- Tobin, J., 1958, “Estimation of relationships for limited dependent variables”, en *Econometrica*. 26 (1): 24-36.

- Urrea-Giraldo, F. y Rodríguez-Sánchez, D.A., 2014, “Collecting ethnic and racial data in censuses and surveys: the latin american experience in the cases of Brazil, Colombia, Mexico and Peru”, en *Revista de Historia Comparada*, 8(1), 7-35.
- Valdés, L.M., 2001, “Los indios del tercer milenio”, en *Ciencias*, (60-61) Octubre 2000-Marzo 2001.
- Vázquez-Sandrín, G. y Quezada, M.F., 2015, “Los indígenas autoadscritos de México en el censo 2010: ¿Revitalización étnica o sobreestimación censal?”, en *Papeles de Población*, 21(86), 171-218.
- Wade, P., 1997, *Race and ethnicity in Latin America*, Chicago, Pluto Press.
- Wade, P., 2008, “Race in Latin America”, en Poole, Deborah, *A companion to Latin American Anthropology*, Wiley-Blackwell Publishing Ltd., Boston, págs. 177-192.
- Wooldridge, J.M., 2002, *Econometric Analysis of Cross Section and Panel Data*. The MIT Press, Cambridge, Massachusetts. London, England.

## **RESUMEN CURRICULAR DE LOS AUTORES**

### *Ignacio Ibarra López*

Es Director de Carrera de Economía, Campus Puebla del Tecnológico de Monterrey. Director de la Carrera de Economía, Departamento de Economía de la Escuela de Ciencias Sociales y Gobierno del Tecnológico de Monterrey, Campus Ciudad de Puebla. Especialista en evaluación de programas sociales y finanzas públicas. Ha participado como asesor y consultor de la Comisión Centroamericana de Estadísticas del Sistema de Integración Centroamericana (SICA), el Banco Interamericano de Desarrollo y la Comisión Nacional Bancaria y de Valores. Realizó el proyecto financiado por CONACYT “Estimación Multidimensional de la Población Indígena en México” con el cual se busca identificar las razones por las cuales está subestimada la población indígena en el país.

Dirección electrónica: [iibarral@tec.mx](mailto:iibarral@tec.mx)

Registro ORCID: <https://orcid.org/0000-0002-0371-3976>

### *Hector Calleros*

Profesor adjunto del Centro de Estudios Americanos de la Universidad de Varsovia (Ośrodek Studiów Amerykańskich, Uniwersytet Warszawski) en Polonia. Tiene un doctorado en Ciencias Políticas de la Universidad de Leeds y una maestría en Estudios Legislativos de la Universidad de Hull en Reino Unido; obtuvo su Licenciatura en Ciencias Políticas, en la Universidad Nacional Autónoma de México (UNAM). Su interés se centra en el estudio de las legislaturas, movimientos sociales y etnicidad. Ha realizado

estancias de investigación en las universidades de Varsovia y Bucarest. Previamente fue profesor e investigador en COLTLAX. Es miembro tanto de la Asociación Internacional de Sociología (ISA) y de la Asociación Internacional de Ciencias Políticas (IPSA).

Dirección electrónica: [iibarral@tec.mx](mailto:iibarral@tec.mx)

*Jorge David Cortés Moreno*

Actualmente se encarga de la Red Nacional de Librerías en el Fondo de Cultura Económica y Educal. Tiene un Posdoctorado en Administración, Educación, Políticas Públicas y Gobierno por The New Mexico University, Estados Unidos. Doctor en Administración Pública por el Instituto de Administración Pública, de Puebla. Maestro en Comunicación Política y Gobernanza, también por The George Washington University. Maestro en Ciencias Sociales por El Colegio Mexiquense. Licenciado en Administración Pública por la BUAP. Es autor de diversos artículos académicos y coautor de más de siete libros sobre temas de Ciencia Política.

Dirección electrónica: [iibarral@tec.mx](mailto:iibarral@tec.mx)

Artículo recibido el 13 de noviembre de 2018 y aprobado el 18 de diciembre de 2018.