



*Conjunto de herramientas para
medir el comercio ilícito de
los productos de tabaco*

Cita sugerida: Stoklosa, M., Paraje, G., Blecher, E., Conjunto de herramientas para medir el comercio ilícito de los productos de tabaco. Un Conjunto de herramientas de Tobacconomics y la Sociedad Americana del Cáncer. Chicago, IL: Tobacconomics, Health Policy Center, Institute for Health Research and Policy, University of Illinois at Chicago, 2020. www.tobacconomics.org

Autores: Este Conjunto de herramientas fue escrito por el Dr. Michal Stoklosa de la Sociedad Americana del Cáncer en Atlanta, Estados Unidos; el Dr. Guillermo Paraje, profesor de la Escuela de Negocios de la Universidad Adolfo Ibáñez, Santiago de Chile; y el Dr. Evan Blecher, economista sénior del Centro de Políticas de Salud, University of Illinois at Chicago. La revisión por pares fue proporcionada por el Dr. Roberto Iglesias, oficial técnico de Políticas Fiscales para la Salud de la Organización Mundial de la Salud, Ginebra, Suiza; y la Dra. Hana Ross, oficial principal de investigación de la Unidad de Investigación de Economía de los Productos Sujetos a Impuestos al Consumo, University of Cape Town, Ciudad del Cabo, Sudáfrica.

Este Conjunto de herramientas fue financiado por Bloomberg Philanthropies.

Sobre Tobacconomics: Tobacconomics es el resultado de la colaboración de destacados investigadores que desde hace casi treinta años estudian los aspectos económicos de las políticas de lucha contra el tabaco. El equipo se dedica a facilitar a investigadores, defensores y responsables políticos el acceso a los mejores y más recientes trabajos de investigación sobre qué funciona, o no funciona, a la hora de reducir el consumo de tabaco y sus repercusiones económicas. Como programa de la University of Illinois at Chicago, Tobacconomics no está vinculado a ningún fabricante de tabaco. Visite www.tobacconomics.org o síganos en Twitter www.twitter.com/tobacconomics.

Acerca de Investigación de Políticas Económicas de la Salud (EHPR, por sus siglas en inglés): El programa EHPR de la Sociedad Americana del Cáncer busca luchar contra el cáncer en todo el mundo a través de la investigación de los aspectos económicos y políticos de los factores de riesgo del cáncer. Estos factores de riesgo incluyen las áreas que abarca el tabaco, la alimentación saludable y una vida activa (HEAL, por sus siglas en inglés), y el consumo dañino del alcohol. En el ámbito de la lucha contra el tabaco, el equipo ha desarrollado conocimiento líder en el mundo en temas económicos, en particular en lo que respecta a las políticas fiscales como los impuestos, el comercio y la inversión. El equipo también ha liderado la producción de las recientes ediciones de *El Atlas del Tabaco*, uno de los recursos más influyentes y citados en la lucha contra el tabaco. Visite el sitio web complementario del libro: tobaccoatlas.org.

Mejorando nuestro Conjunto de herramientas: Los equipos de Tobacconomics y EHPR están comprometidos a lograr que este Conjunto de herramientas sea lo más claro y útil posible. Nos gustaría conocer su opinión sobre si este Conjunto de herramientas le resultó útil en su investigación y, de ser así, le agradeceríamos que nos compartiera su experiencia sobre cualquier implementación exitosa que haya tenido. También nos gustaría saber si ha encontrado algún problema al aplicar las metodologías que le presentamos en este Conjunto de herramientas, así como sus opiniones sobre cómo podríamos mejorarlo.

Para cualquier comentario o preguntas sobre este Conjunto de herramientas y su contenido, escribanos un correo electrónico a info@tobacconomics.org. Nos encantaría escucharlo.

Agradecimientos

Los autores desean agradecer al Dr. Jeffrey Drope, a la Dra. Liora Sahar y al Dr. Robert Stephens de la Sociedad Americana del Cáncer por sus aportaciones en cuanto a los métodos de recopilación de datos primarios.

Tabla de contenido

1	<i>Introducción</i>	3
1.1	Propósito	5
1.2	Quién debería usar este Conjunto de herramientas	5
1.3	Cómo utilizar este Conjunto de herramientas	5
2	<i>Algunas descripciones y definiciones</i>	7
3	<i>Estudios sobre el análisis de cajetillas</i>	9
3.1	Introducción	9
3.2	Planeación de la investigación	10
3.2.1	Formulación de preguntas de investigación	10
3.2.2	Selección de las características de las cajetillas para su análisis	11
3.2.3	Elección del método de investigación	14
3.2.3.1	Ventajas de la encuesta sobre el análisis de cajetillas de fumadores	15
3.2.3.2	Ventajas del método de recolección de cajetillas desechadas	16
3.2.3.3	Otros métodos de análisis de cajetillas	17
3.3	Muestreo	19
3.3.1	Población de estudio	19
3.3.2	Geografía	20
3.3.3	Edad y otras características de las personas	21
3.3.4	Tamaño de la muestra	21
3.3.5	Diseño de la muestra	23
3.3.5.1	Muestreo aleatorio simple	23
3.3.5.2	Muestreo estratificado	24
3.3.5.3	Muestreo por conglomerado	25
3.3.6	Etapas de muestreo	29
3.4	Recopilación de datos	31
3.4.1	Construcción del cuestionario	31
3.4.2	Aprobación ética	34
3.4.3	Ejecución de la encuesta	34
3.4.4	Identificación de cajetillas falsificadas	36
3.4.4.1	Creación y calibración de la prueba	37
3.4.4.2	Determinar el alcance de la falsificación de cigarrillos mediante la prueba	38
3.4.5	Uso de estudios de análisis de cajetillas para otros productos de tabaco	38
3.5	Análisis de datos	38
3.6	Conclusión	39

4.	<i>Modelo de brecha</i>	40
4.1	Introducción	40
4.2	Marco conceptual de un modelo de brecha	40
4.3	Fortalezas, debilidades y limitaciones	42
4.4	Fuentes de datos	44
4.4.1	Estimación de las ventas con impuestos pagados utilizando datos oficiales	44
4.4.2	Estimación de las ventas con impuestos pagados utilizando fuentes externas	44
4.4.3	Encuestas sobre la prevalencia del tabaquismo	45
4.4.4	Datos de población	47
4.5	Desafíos con los datos	47
4.5.1	Anticipando cambios fiscales	47
4.5.2	Inconsistencia en los períodos de tiempo	48
4.5.3	Subregistro de la prevalencia e intensidad del tabaquismo	49
4.5.4	Sesgo del periodo recordatorio de la intensidad del tabaquismo	49
4.6	Estimación de la brecha y presentación de los datos	50
4.7	Casos prácticos	54
4.7.1	El caso del Canadá	54
4.7.2	Análisis de brechas para cinco países de América Latina	55
4.7.3	Estimación del comercio ilícito en Sudáfrica	55
4.7.4	Medición de las brechas fiscales en el Reino Unido	56
4.8	Conclusión	56
5.	<i>Bibliografía</i>	58
6.	<i>Anexos de Metodología y Código</i>	63
6.1	Selección de puntos de congregación de alto tráfico en Santiago de Chile	63
6.2	Instrumento de encuesta de muestra	64
6.3	Muestra de solicitud de exención presentada a la Junta de Revisión Institucional	71
6.4	Muestreo de regiones geográficas utilizando el método de probabilidad proporcional al tamaño (PPT) de prevalencia de tabaquismo prevista en Brasil	72
6.5	Muestreo aleatorio simple	74
6.6	Ponderación multivariante: ejemplo de Chicago	74
6.7	Imputar el estado de cumplimiento para fumadores que no mostraron cajetillas a los encuestadores: enfoque de regresión logística	75

Introducción

1

La forma más eficiente de reducir el consumo de tabaco es aumentando el precio de los productos de tabaco a través de impuestos más altos (IARC 2011 y NCI/OMS 2016). El aumento de los precios reduce el consumo de tabaco de múltiples maneras: mayor cesación tabáquica; reducción de la intensidad de uso de tabaco entre los consumidores habituales; y menor iniciación en el consumo de tabaco, especialmente entre los jóvenes. Los aumentos en los impuestos al tabaco también generan ingresos gubernamentales adicionales. Sin embargo, la industria tabacalera suele oponerse a los aumentos en los impuestos al tabaco utilizando el argumento principal de que los impuestos son ineficientes para alcanzar los objetivos de las políticas del gobierno debido a los aumentos en el comercio ilícito. La industria tabacalera también utiliza la amenaza del creciente comercio ilícito como estrategia para oponerse a otras medidas de la lucha contra el tabaco; incluidas las prohibiciones sobre los cigarrillos mentolados, el empaquetado simple, las leyes sobre el tamaño mínimo de las cajetillas y las prohibiciones de exhibición en los puntos de venta.

En vista de que la industria tabacalera utiliza el comercio ilícito para oponerse a los aumentos de los impuestos al tabaco, es importante entender el verdadero alcance y la naturaleza del comercio ilícito de tabaco. Este entendimiento es crucial para formular nuevas regulaciones contra el tabaco. El comercio ilícito hace que los productos de tabaco sean más asequibles y accesibles. El aumento de la asequibilidad y la accesibilidad da lugar a un mayor consumo total de tabaco, lo que se traduce a una mayor mortalidad y morbilidad entre los consumidores. Esto, a su vez, obstaculiza el crecimiento económico por el aumento de los costos de la atención médica y la reducción de la productividad en los trabajadores. Además, el comercio ilícito de productos de tabaco aumenta las tasas de delincuencia, socava la recaudación de impuestos y degrada los sistemas administrativos, así como las políticas fiscales.

En 2012, las Partes que conforman el Convenio Marco de la Organización Mundial de la Salud (OMS) para el Control del Tabaco (CMCT) abordaron la cuestión del comercio ilícito de productos de tabaco a nivel mundial mediante la adopción del Protocolo para la Eliminación del Comercio Ilícito de Productos de Tabaco. Este nuevo tratado entró en vigor en 2018. El Protocolo proporciona herramientas para prevenir el comercio ilícito asegurando la cadena de suministro (incluida la implementación de sistemas de rastreo y localización), exigiendo medidas estrictas de aplicación de la ley y fortaleciendo la cooperación internacional.

Junto con este Conjunto de herramientas, Tobacconomics ha publicado un reporte titulado “El comercio ilícito de productos de tabaco no debe obstaculizar las reformas ni los aumentos de la política fiscal sobre el tabaco”. El reporte, junto con el informe sobre políticas públicas y las notas descriptivas de los casos prácticos de países que lo acompañan, están diseñados para informar a los investigadores, los profesionales y a los encargados de la formulación de políticas públicas de control del tabaco y proporcionar una mejor comprensión del comercio ilícito de los productos de tabaco. En los documentos se evalúa críticamente la oposición que encabeza la industria tabacalera contra las reformas fiscales a los productos de tabaco y se llega a la conclusión de que las reformas de la política fiscal del tabaco no deben desalentarse con el argumento de la industria de que se da un aumento en el comercio ilícito.

Mensajes clave del Reporte “El comercio ilícito de productos de tabaco no debe obstaculizar las reformas ni los aumentos de la política fiscal sobre el tabaco”:



1. La industria tabacalera utiliza el comercio ilícito para oponerse a los aumentos de los impuestos, argumentando específicamente que los aumentos de los impuestos al tabaco darán lugar a incrementos en el comercio ilícito, lo que a su vez socavará los objetivos de salud pública y de política fiscal.



2. La industria tabacalera exagera la escala y el alcance del comercio ilícito como una forma de abogar contra el aumento de los impuestos al tabaco.



3. Los impuestos y los precios no son el principal impulsor ni el factor determinante del comercio ilícito; es más probable que muchos otros factores impulsen el comercio ilícito.



4. Incluso en presencia del comercio ilícito, la experiencia de muchos países demuestra que los aumentos a los impuestos al tabaco han producido sistemáticamente beneficios fiscales y sanitarios importantes gracias al aumento de los ingresos y la reducción del consumo de tabaco.



5. Si a los gobiernos les preocupan los niveles y/o el alcance del comercio ilícito, hay muchas políticas y medidas administrativas y de aplicación de la ley que pueden adoptar para reducir el comercio ilícito, incluso mientras aumentan los impuestos al tabaco.

Dada la explotación del comercio ilícito por parte de la industria tabacalera en el espacio de la política fiscal del tabaco, ha aumentado la necesidad de contar con estimaciones confiables del nivel y las tendencias del comercio ilícito. El comercio ilícito de tabaco es difícil de estimar, ya que las actividades subyacentes son ilegales y a menudo se realizan en un mercado negro que es inobservable. Los vendedores ocultan sus productos ilícitos, mientras que los usuarios ocultan sus hábitos de fumar ilegalmente. Debido a la naturaleza inobservable del comercio ilícito, a menudo es necesario realizar investigaciones originales, incluyendo la recopilación de datos primarios.

Pocos gobiernos disponen de estimaciones sobre el comercio ilícito de tabaco. La mayoría de los gobiernos no realizan ni encargan a otros sus propios estudios sobre el comercio ilícito. En cambio, la industria tabacalera les proporciona a los gobiernos estimaciones sobre el comercio ilícito de tabaco. No es de sorprenderse que, en la mayoría de los casos, los estudios financiados por la industria tabacalera exageren sistemática y estratégicamente el comercio ilícito. Por lo tanto, es necesario realizar estudios confiables, transparentes, reproducibles e independientes. Las estimaciones independientes o académicas que utilizan métodos confiables están sujetas a las revisiones por pares y, lo más importante, no están financiadas por la industria tabacalera. Este tipo de investigación es fundamental para estimar el nivel y las tendencias del comercio ilícito para poder contrastar los resultados sesgados de la industria tabacalera y proporcionarle al público, a la sociedad civil y al gobierno, una comprensión más precisa del comercio ilícito.

Se han desarrollado varios métodos, utilizando tanto datos primarios como secundarios, para estimar el nivel y/o las tendencias del comercio ilícito. En una publicación anterior de *Tobacconomics*, Ross (2015) proporcionó un examen exhaustivo de los métodos existentes, incluyendo sus ventajas y desventajas. En vez de reproducir el examen de Ross, este conjunto de herramientas ofrece una orientación técnica detallada para estimar el comercio ilícito de los productos de tabaco a través de los dos enfoques más utilizados. Esta orientación incluye el examen de cuestiones relativas a la recopilación de datos primarios, el análisis de datos y las técnicas econométricas relevantes con códigos estadísticos. También se esbozan elementos de planeación de la investigación, diseño de encuestas, muestreo e identificación de cajetillas.

1.1 Propósito

El objetivo principal de este conjunto de herramientas es guiar las investigaciones sobre el comercio ilícito de tabaco. El conjunto de herramientas proporciona una guía técnica detallada sobre la aplicación de los dos enfoques más utilizados para estimar el comercio ilícito: en primer lugar, la estimación del comercio ilícito utilizando la recopilación de datos primarios mediante la recolección de cajetillas físicas y encuestas a fumadores; y en segundo lugar, la estimación del comercio ilícito utilizando datos secundarios mediante un modelo de brechas. En este conjunto de herramientas se examinan las herramientas y técnicas estadísticas y económicas que pueden utilizarse para analizar los datos así como los antecedentes teóricos y los fundamentos de cada método. El conjunto de herramientas también propone métodos de estimación y uso del software estadístico Stata®.

Este conjunto de herramientas está especialmente orientado a la investigación en países de ingreso mediano y bajo en los que tal vez no existan estimaciones independientes sobre el comercio ilícito. En muchos países, los datos secundarios, recopilados como parte de la prevalencia de tabaquismo o de las encuestas de hogares, se pueden utilizar para estimar el tamaño del mercado ilícito de tabaco. Asimismo se dispone de una recopilación de datos primarios rentable. Los datos primarios y secundarios permiten una evaluación rentable, confiable e independiente de la magnitud del problema.

Este conjunto de herramientas es uno de varios desarrollados por el Banco Mundial, la Organización Mundial de la Salud y Tobacconomics para ofrecer una guía para realizar análisis económicos de la demanda de tabaco y de los impactos del consumo de tabaco en el empleo, en la equidad, en el comercio ilícito y en los costos económicos. Este conjunto de herramientas de Tobacconomics es el segundo de la serie de Tobacconomics diseñados para desarrollar capacidad de análisis económico de los impuestos al tabaco. El primer conjunto de herramientas se publicó en 2019 y se centró en la utilización de encuestas de gasto de los hogares para estimar la demanda de productos de tabaco (John *et ál.*, 2019).

1.2 Quién debería usar este Conjunto de herramientas

El conjunto de instrumentos no presupone que el lector tenga conocimientos sobre los impuestos al tabaco, la economía del control del tabaco o los métodos de diseño de encuestas. Sin embargo, se requiere una formación en economía, econometría/estadística y una comprensión básica de Stata. Armados con los conocimientos que proporciona este conjunto de herramientas, los investigadores estarán preparados para participar en estudios independientes sobre el comercio ilícito de productos de tabaco a través de la discusión de métodos y guías paso a paso con Stata. Los responsables de la formulación de políticas, los analistas gubernamentales y las organizaciones de la sociedad civil también se pueden beneficiar de los debates sobre políticas, la justificación de los diferentes conceptos económicos en la lucha contra el tabaco y las interpretaciones de los resultados proporcionados. Por último, este conjunto de herramientas puede ayudarles a todos los usuarios a comprender mejor la estimación del alcance del comercio ilícito de tabaco, realizar una estimación independiente del comercio ilícito e identificar y comprender los problemas que plantean las estimaciones proporcionadas por la industria tabacalera.

1.3 Cómo utilizar este Conjunto de herramientas

Los capítulos están estructurados de la siguiente manera: cada capítulo comienza con una introducción y una explicación de los principios fundamentales, junto con la justificación de la participación en el análisis. A continuación, se realiza una breve discusión técnica sobre la recopilación de datos y los métodos econométricos. La discusión de los métodos econométricos se mantiene intencionalmente al mínimo, ya que la mayoría de los detalles están disponibles en los libros de texto comunes de econometría y otras fuentes publicadas. Se ofrecen referencias adicionales en caso de que el lector desee buscar información adicional sobre estos conceptos teóricos. Después de la sección de métodos, se presenta una breve discusión sobre la preparación de los datos para el análisis. La discusión incluye el uso de Stata como una herramienta de análisis así como los códigos relevantes de Stata. Se presentan casos prácticos de varios países. Estos casos prácticos demuestran las variaciones metodológicas que a menudo se requieren para evaluar el tamaño del mercado ilícito de tabaco. En el conjunto de herramientas se examinan los métodos de análisis relevantes para todos los productos de tabaco, pero los casos prácticos se centran en los cigarrillos.

También cabe señalar que no existe un enfoque único para las estimaciones del comercio ilícito. Cada país tendrá sus propias y exclusivas circunstancias en lo que respecta la disponibilidad de datos, la composición del mercado de tabaco, etc. Por lo tanto, además de obtener conocimientos e inspiración en este conjunto de herramientas, se les invita a los investigadores que se inician en la medición del comercio ilícito de cigarrillos a que pidan asesoramiento a investigadores más experimentados. Este conjunto de herramientas ofrece una visión general de las investigaciones realizadas por algunos de los investigadores más experimentados en este campo.

El conjunto de herramientas está organizado de la siguiente manera: en el Capítulo 2 se ofrece un análisis de antecedentes sobre las definiciones de comercio ilícito. El comercio ilícito abarca tanto la elusión como la evasión fiscal, y dentro del ámbito de la elusión y la evasión fiscal, hay diferentes tipos. Metodologías específicas pueden ser más apropiadas que otras para estimar los diferentes tipos de elusión y evasión fiscal. Por consiguiente, es imperativo para la aplicación del conjunto de herramientas que los lectores comiencen con una comprensión fundamental de los diferentes tipos y atributos del comercio ilícito.

En el Capítulo 3 se examina la utilización de la recopilación de datos primarios para estimar el comercio ilícito de tabaco. Se centra en dos métodos. El primer método consiste en la inspección de las cajetillas de cigarrillos a través de la recolección de estas. También se describe la recopilación de datos de los vendedores de cigarrillos individuales. El segundo método describe el método de encuestas de análisis de las cajetillas de cigarrillos de fumadores. Estos dos métodos ofrecen las estimaciones más confiables sobre el comercio ilícito de tabaco por dos razones. Primero, los dos métodos pueden medir el alcance y las tendencias del comercio ilícito. Segundo, pueden determinar la naturaleza y el origen del comercio ilícito, así como, en el caso de la encuesta a fumadores, las características de los consumidores que adquieren esos productos.

En el Capítulo 4 se examina la utilización de datos secundarios para estimar el comercio ilícito de tabaco a través el modelo de brechas. El modelo de brechas suele ser una forma rápida y rentable de estimar la tendencia y, a veces, el alcance del comercio ilícito. Sin embargo, este método plantea varios problemas y cada uno de ellos se describe detalladamente en el capítulo, junto con sugerencias sobre la forma de evitar esos problemas. Cuando la recopilación de datos primarios se aborda correctamente mediante validación cruzada, el método de análisis de brechas es una forma confiable y precisa de estimar el comercio ilícito de tabaco. En este capítulo también se incluye un código de Stata paso a paso para estimar el comercio ilícito utilizando el método de análisis de brechas.

El Capítulo 4 y los Anexos de Metodología y Código (Capítulo 6) muestran ejemplos específicos de los métodos de estudio para estimar el comercio ilícito de productos de tabaco. El Capítulo 4 ofrece un código de Stata paso a paso para estimar el comercio ilícito utilizando el modelo de brechas, mientras que el Capítulo 6 ofrece ejemplos prácticos y un código de Stata paso a paso para mostrar las técnicas apropiadas para estimar el comercio ilícito utilizando el enfoque de análisis de cajetillas. Las secciones de código en los dos capítulos se basan en los antecedentes de Stata que se analizan en el anexo del conjunto de herramientas anterior (John *et ál.*, 2019). Los comandos individuales de Stata se colocan entre corchetes angulares `< >` y están en cursiva. Esto es solo para fines ilustrativos. El comando en sí se debe usar sin corchetes para que el código se ejecute en Stata. Los nombres de las variables utilizadas en los ejemplos están en cursiva.

Algunas descripciones y definiciones

2

Las definiciones inconsistentes de comercio ilícito crean confusión, lo que puede llevar a los profesionales o a los responsables de la creación de políticas públicas a sacar conclusiones incorrectas y afectar negativamente los resultados de las políticas. El Protocolo para la Eliminación del Comercio Ilícito de Productos de Tabaco del CMCT de la OMS (OMS, 2013), define al comercio ilícito como “toda práctica o conducta prohibida por la ley, relativa a la producción, envío, recepción, posesión, distribución, venta o compra, incluida toda práctica o conducta destinada a facilitar esa actividad”. A partir de la amplia definición de comercio ilícito de la OMS, este conjunto de herramientas se centra en el comercio ilícito en lo que respecta al comercio con el pago parcial o sin el pago total de los impuestos.

Sin embargo, antes de examinar las definiciones de comercio ilícito, puede ser útil definir aquello que no se considera comercio ilícito de productos de tabaco:

- La producción nacional por la que se pagan todos los impuestos nacionales (por ejemplo, el IVA, el impuesto general a las ventas, los impuestos al tabaco, etc.).
- La producción extranjera que se importa legalmente al mercado nacional con derechos de importación totalmente pagados (si así se exige) y con impuestos nacionales.
- Compras libres de impuestos dentro de la asignación legal y que se usan solo para consumo personal (no para ser revendidas).

Por el contrario, los métodos ilegales para evitar los impuestos al tabaco se denominan evasión fiscal, ya que tienen por objeto evadir el pago de algunos o todos los impuestos al tabaco. La evasión fiscal suele ocurrir debido a una aplicación deficiente de la ley. Hay varios tipos de evasión fiscal. Una de las formas más comunes es el contrabando de productos de tabaco a través de fronteras sin pagar impuestos en la jurisdicción de consumo. Esto también puede ocurrir cuando las mercancías se desvían durante el tránsito. En muchos casos, es posible que se hayan pagado impuestos en otra jurisdicción, aunque sea una con impuestos más bajos. Aunque el contrabando transfronterizo es la forma más común de comercio ilícito a nivel mundial, la evasión fiscal también ocurre con la producción nacional.

Los cigarrillos falsificados son cigarrillos fabricados sin la autorización del propietario de la marca. La intención es engañar a los consumidores sobre el origen de los cigarrillos y evitar el pago de impuestos. Los cigarrillos ilícitos o “cigarrillos de etiqueta blanca baratos” son marcas fabricadas en una jurisdicción, a menudo legalmente en la jurisdicción de fabricación, que luego se pasan por contrabando y se venden en otra jurisdicción en la que no se pagan todos los impuestos aplicables. Los productos de tabaco sin marca se suelen vender como tabaco suelto de corte fino. Los cigarrillos sin marca se venden en bolsas de plástico transparentes y se llaman “baggies”. La fabricación ilícita puede consistir en la distorsión de la calidad y el origen, el no obtener una licencia para cultivar y producir tabaco y/o el no registrarse como importador/exportador/distribuidor.

El comercio ilícito lo realizan entidades que no están debidamente registradas en los organismos gubernamentales o entidades legítimas que operan fuera de los límites legales y reglamentarios. La motivación subyacente para participar en el comercio ilícito de tabaco suele estar vinculada a la magnitud de la operación ilegal. Las operaciones de evasión fiscal en pequeña escala suelen tener lugar entre países vecinos o a nivel

regional. Esto implica mover productos a través de la frontera por encima de los límites permitidos. El contrabando también puede ocurrir cuando los productos comprados “para consumo personal” en un país se venden con fines de lucro en otro sin pagar los impuestos correspondientes (conocido como “contrabando hormiga”). Las operaciones a pequeña escala puede que paguen impuestos a sus productos, pero en una jurisdicción con impuestos menores. El contrabando generalmente ocurre a pequeña escala. Sin embargo, en los casos en que no existen límites diarios para el cruce de fronteras o en que se pueden cruzar las fronteras a bajo costo, el contrabando puede contribuir considerablemente a la magnitud general del comercio ilícito de tabaco. Por consiguiente, al estimar el comercio ilícito, es importante tener en cuenta la evasión fiscal a pequeña escala.

La evasión fiscal a gran escala suele implicar el no pagar todos los impuestos y a menudo se produce a nivel internacional. La motivación de estas acciones es la avaricia individual o corporativa, el lavado de dinero o el financiamiento de actividades criminales. Las operaciones de evasión fiscal a gran escala pueden consistir en falsificaciones, productos auténticos con sellos fiscales falsificados, cigarrillos ilícitos o producción nacional mayor a las cantidades declaradas. A menudo se aprovechan de los regímenes de “en tránsito” y/o de las zonas libres de impuestos (Ross, 2015).

La elusión fiscal se diferencia de la evasión fiscal en que se refiere a los mecanismos legales para evitar el pago de impuestos. La elusión fiscal se aprovecha de una política o administración deficiente. Una forma destacada de elusión fiscal de las personas es la compra transfronteriza. Los consumidores, a menudo consumidores individuales de tabaco, compran productos de tabaco en una jurisdicción con impuestos más bajos (o libres de impuestos) dentro de la cantidad permitida.

Las empresas tabacaleras aprovechan las lagunas de la legislación mediante prácticas contables inteligentes para reducir su responsabilidad fiscal (Ross *et ál.*, 2017). Los fabricantes de cigarrillos utilizan la técnica de anticipación, práctica que consiste en producir un número significativamente mayor de cigarrillos antes de que entre en vigor un aumento de impuestos. De esta manera, los fabricantes evitan pagar un porcentaje de impuesto mayor en el futuro (Ross *et ál.*, 2017). Los fabricantes también pueden cambiar los atributos de los productos para evitar un aumento de impuestos. Estos cambios incluyen la reducción del peso del tabaco presente en un cigarrillo o el cambio de varias características para desplazar sus productos a una categoría fiscal más baja.

Estos tipos de actividades de elusión fiscal son técnicamente legales. Sin embargo, privan al gobierno de los ingresos fiscales y contribuyen a la asequibilidad de los productos de tabaco. Por lo tanto, estas actividades de elusión fiscal socavan las políticas de salud pública y las medidas fiscales. Los gobiernos deberían prestar especial atención a los planes de elusión fiscal a gran escala, ya que son los que más ingresos dejan de percibir y los que más contribuyen a la carga de la salud pública. En muchos casos, el gobierno puede abordar directamente el problema de la elusión fiscal mediante la supervisión administrativa o mejorando el diseño de las políticas (Ross *et ál.*, 2017). Tanto la OMS (2010) como el Banco Mundial (2018) han publicado guías útiles para ayudar a los países a abordar el problema de la elusión fiscal.

Aunque la elusión y la evasión fiscal funcionan de manera diferente, a menudo se explotan con el mismo fin. Ambas se benefician de la reducción en la responsabilidad fiscal y socavan los objetivos de salud pública y de políticas fiscales de los impuestos al tabaco. A lo largo de este conjunto de herramientas se proporcionan descripciones detalladas de cada uno de los tipos específicos de evasión o elusión fiscal que se están midiendo.

Estudios sobre el análisis de cajetillas

3

3.1 Introducción

Como se ha mencionado anteriormente, es difícil obtener una medida directa del comercio ilícito de tabaco debido al secreto inherente en torno a la compra y el uso de productos de tabaco ilícitos. Sin embargo, los investigadores pueden evitar este problema de dos maneras. La primera es observando directamente el empaquetado del producto. De forma alternativa, los investigadores pueden preguntarles a los fumadores sobre las características de su producto, en lugar de preguntarles directamente sobre el consumo de cigarrillos ilícitos. El empaquetado legal e ilegal de los cigarrillos a menudo difiere por características observables, entre ellas la marca de la cajetilla, la presencia de un sello fiscal adecuado, la presencia de advertencias sanitarias adecuadas, la falta de información sobre el precio o el contenido y los marcados de la venta libre de impuestos. Los investigadores pueden recopilar información sobre las cajetillas de cigarrillos evaluando las cajetillas que han sido desechadas o las cajetillas en posesión de los fumadores.

El consumo humano y otros comportamientos son analizados a través de la basura en diferentes ciencias sociales. La basura revela mucho sobre el consumo, incluido el consumo ilícito. Así pues, el consumo ilícito de tabaco se observa fácilmente mediante el análisis de la basura. Las cajetillas desechadas se sacan directamente de las calles o de la basura. Del mismo modo, los investigadores pueden observar los tipos de cajetillas que tienen los fumadores, ya sea por observación aleatoria pública o privada en el hogar del fumador. Debido al estigma asociado al consumo de cigarrillos ilícitos, es poco probable que los consumidores admitan de buena fe que consumen cigarrillos ilícitos. Sin embargo, no lo piensan dos veces a la hora de indicar la marca y el precio de sus cigarrillos o de mostrar la cajetilla. Esta información puede proporcionarles a los investigadores una valiosa información sobre la prevalencia de las cajetillas que no cumplen con los impuestos.

El presente capítulo se centra en dos enfoques del análisis de cajetillas: las cajetillas desechadas y las encuestas a fumadores. Este ofrece una guía paso a paso a los investigadores que desean aplicar estos métodos para estimar la prevalencia de los productos de cigarrillos que no cumplen con las regulaciones. Se proporciona una referencia a los estudios que han utilizado una o ambas metodologías. El capítulo se refiere principalmente a la estimación de la falta de cumplimiento fiscal de los cigarrillos. Los cigarrillos son el producto de tabaco más común y, por lo tanto, son de gran interés para los responsables de formular políticas públicas. Sin embargo, las estrategias utilizadas para los cigarrillos también se pueden aplicar para estimar el alcance del comercio ilícito de otros productos de tabaco (véase la Sección 3.4.5).

El capítulo analiza cada etapa de la investigación. En primer lugar, describe las cuestiones de planeación de la investigación, incluida la formulación de la pregunta de investigación y la selección de las características de la cajetilla para su análisis. Después se examinan brevemente las ventajas y desventajas de los métodos de recolección de cajetillas desechadas y del análisis de cajetillas en posesión de los fumadores para ayudar a los investigadores a elegir el método adecuado para su proyecto de investigación. A continuación, se explican detalladamente las cuestiones relativas a la selección de muestras, ya que hay muchos aspectos en los que el muestreo de las cajetillas difiere del muestreo de las personas. Por último, en el capítulo se examinan cuestiones relativas a la elaboración de cuestionarios, la recopilación de datos y el análisis.

3.2 Planeación de la investigación

Los investigadores tienden a centrar su atención en el diseño y análisis de los estudios. Si bien estos elementos de investigación son sin duda de gran importancia, la recopilación de datos merece también una planeación cuidadosa. Especialmente cuando se trata de medir y estimar el alcance del comercio ilícito, es esencial recopilar cuidadosamente los datos para poder realizar estimaciones precisas. La naturaleza encubierta del comercio ilícito hace que las estimaciones precisas sean aún más valiosas para los responsables de la formulación de políticas a medida que elaboran políticas públicas para frenar el comercio ilícito. Además, si el objetivo es ayudar a los responsables de la formulación de políticas a crear políticas públicas efectivas, entonces la investigación debería centrarse en cuestiones de importancia primordial para los responsables de la formulación de políticas. También es necesario tener en cuenta las limitaciones de tiempo y financiamiento en el proceso de investigación y recopilación de datos. Un plan de investigación bien diseñado debería equilibrar todas las necesidades de investigación y asegurar un diseño y análisis cuidadoso de los estudios, así como estimaciones valiosas y precisas para los responsables de formular las políticas. Estos problemas se analizan más adelante.

3.2.1 Formulación de preguntas de investigación

Una comprensión fundamental del tabaquismo y del comercio ilícito de tabaco es un antecedente esencial necesario para formular una pregunta y una hipótesis de investigación relevante. Es imperativo que los investigadores se informen antes de emprender una investigación. Se puede obtener una comprensión básica de estas cuestiones examinando los informes de los medios de comunicación, la bibliografía académica y las políticas gubernamentales relacionadas con el comercio ilícito de tabaco. La información clave para mejorar el diseño del estudio incluye la comprensión del comportamiento del consumidor en lo que se refiere a la compra y el consumo de cigarrillos, tanto legales como ilegales, así como la comprensión de los orígenes de los cigarrillos ilícitos. Además, es necesario comprender el público al que se dirige la investigación, ya sean los encargados de formular políticas, los profesionales, los investigadores u otros actores, y ajustar las preguntas de investigación en función del público.

La información que se recopila durante el proceso de preparación determina en gran medida la pregunta y el método de investigación. Algunos ejemplos de circunstancias importantes que deben considerarse en la fase de preparación son: si una gran proporción de los cigarrillos vendidos es en forma de cigarrillos sueltos o individuales, podría ser necesario inspeccionar las cajetillas de los vendedores de estos cigarrillos. Si los cigarrillos se venden principalmente en cajetillas blandas que se degradan fácilmente cuando se desechan, entonces sería preferible realizar una encuesta entre los fumadores en lugar de una inspección de las cajetillas desechadas. Si no es seguro para los encuestadores entrar en ciertas zonas, entonces la encuesta de fumadores en lugares públicos podría ser preferible a las encuestas en el hogar.

Durante esta fase de preparación, es necesario considerar los posibles sesgos. Esto incluye tener en cuenta la posibilidad de sesgos de información que pueden introducirse fácilmente en los estudios. A este respecto, es importante señalar que la industria tabacalera es la principal proveedora de información sobre el comercio ilícito de tabaco. La validación cruzada de esta información muestra que los resultados de la industria tabacalera suelen ser sesgados (Gilmore *et ál.*, 2015). Estos estudios tampoco suelen ser verificados independientemente o revisados por pares (Gallagher *et ál.*, 2019). Además, la industria tabacalera sobrestima o falsifica la amenaza del comercio ilícito con el propósito estratégico de detener la aplicación de nuevas medidas de control del tabaco (Gilmore *et ál.*, 2015). Por lo tanto, la información proporcionada por la industria debe tomarse con precaución debido al conflicto de intereses. Los medios de comunicación deben ser especialmente cautelosos, ya que su público es de gran alcance y la información proporcionada no solo influye en la opinión pública, sino que también tiene repercusiones en la salud pública.

No obstante, es necesario tener en cuenta los informes proporcionados por la industria. Cuando se desarrolla un nuevo diseño de estudio, la investigación preexistente sobre el tabaco es un factor clave a incorporar (por ejemplo, una muestra sesgada). Dado que los informes de la industria tabacalera suelen ser los únicos disponibles sobre el comercio ilícito de tabaco, se convierten en una referencia. Estos informes son utilizados por el gobierno, los medios de comunicación y el público en general. Las investigaciones académicas sobre el comercio ilícito de

tabaco deben identificar los principales defectos de los estudios de la industria tabacalera. De esta manera, los resultados de la industria tabacalera pueden ser abordados y corregidos adecuadamente.

La información proporcionada por el gobierno también puede estar incompleta. Por ejemplo, los controles aduaneros suelen tener una visión específica del comercio ilícito de tabaco, ya que interactúan con el tabaco ilícito que cruza sus fronteras. Según los funcionarios de aduanas, el número de incautaciones en la frontera sería la unidad de medición para la lucha contra el tabaco ilícito. Aunque es una información valiosa, esta medida no logra captar todo el alcance del comercio ilícito de tabaco. La fase de recopilación de información también supone la aclaración de las definiciones relativas al comercio ilícito de productos de tabaco, cuestión que se examinó en el capítulo anterior.

Una vez que se hayan sentado las bases, los investigadores pueden pasar a definir su(s) pregunta(s) de investigación. En este punto es importante considerar el público objetivo de la investigación. Si el estudio está dirigido a los encargados de formular políticas, las preguntas de la investigación se centrarán probablemente en el tamaño del mercado ilícito de cigarrillos o en las posibles repercusiones de las nuevas regulaciones, como los impuestos a los cigarrillos. La mejor manera de ayudar a los responsables de la formulación de políticas es investigar la prevalencia de la elusión y la evasión fiscal. La investigación sobre otros aspectos del comercio ilícito de tabaco sería mejor que se dirigiera a otras partes. Por ejemplo, el origen de una cajetilla de cigarrillos ilícita o las características de los usuarios de cajetillas de cigarrillos ilícitos pueden ser de interés para las entidades que se ocupan de las aduanas y el control de fronteras o los departamentos de salud pública.

Si los investigadores pueden plantear preguntas libremente sin limitaciones de subvenciones o financiamiento, entonces deben hacerlo. Sin embargo, si los recursos son limitados, como suele ser el caso, puede ser útil discutir la(s) cuestión(es) de la investigación con las partes interesadas. Una buena pregunta de investigación debe ser explícita y estar claramente establecida (Kelley *et ál.*, 2003). Sin embargo, las partes interesadas pueden exigir respuestas a una amplia variedad de preguntas, que tal vez no sean viables dentro de las limitaciones de financiamiento, datos o tiempo. Los investigadores deben sacar la mejor pregunta posible del conjunto; de lo contrario, el estudio puede debilitarse. Tal vez no sea factible responder a ciertas preguntas porque los datos no existen o la obtención de estos sería prohibitivamente costosa. Así pues, la experiencia de los investigadores en la materia debería guiar el proceso. En algunos casos, puede bastar con proporcionar una sola estimación nacional del comercio ilícito de cigarrillos para refutar las afirmaciones de la industria.

3.2.2 Selección de las características de las cajetillas para su análisis

Una vez que se establece la pregunta de investigación y se determina el método de recopilación de datos primarios, los investigadores deben considerar en qué características de la cajetilla se deben enfocar. Las características de la cajetilla que ayudan a distinguir entre legales e ilegales pueden incluir la marca de la cajetilla, la presencia de un sello fiscal válido, la advertencia de salud, la información sobre el precio o el contenido, las estampillas de la venta libre de impuestos y otras características. Cada país tendrá un conjunto separado de características de las cajetillas que se utilizarán para distinguir las cajetillas que cumplen y las que no cumplen. Es importante que estas definiciones se creen antes de que comience la recopilación de datos. Durante la fase de preparación, los investigadores deben familiarizarse con las características de la cajetilla que se requieren en su país y crear un protocolo para el proceso del análisis de cajetillas. Como se demuestra más adelante en esta sección, tal vez no sea necesario recopilar datos sobre todas las posibles características de la cajetilla, sino limitar el enfoque a las características que mejor predicen su legalidad.

Además, estas características pueden cambiar con el tiempo. Por esta razón, se aconseja una revisión regular de las leyes que tratan de las características de las cajetillas legales a lo largo del proyecto. Por ejemplo, en el proyecto de ley de control del tabaco de Polonia se establece un modelo de advertencias que los fabricantes de cigarrillos deben colocar en todas las cajetillas. El proyecto de ley también establece lineamientos estrictos en cuanto al tamaño y la colocación de las etiquetas de advertencia (Sejm de la República de Polonia, 2019). Además, el proyecto de ley del impuesto al consumo de Polonia exige que las cajetillas muestren los sellos de dicho impuesto (Sejm de la República de Polonia, 2008).¹ La revisión de estos dos proyectos de ley sería esencial

¹ Las plantillas de los sellos también son proporcionadas por el Ministerio de Hacienda de Polonia (2019).

Gráfico 1 Advertencia sanitaria y sello fiscal requeridos en las cajetillas de cigarrillos en Polonia



Origen: Sejm de la República de Polonia (2019); Ministerio de Hacienda de Polonia (2019).

para la investigación realizada en Polonia, ya que detallan los criterios adecuados para la evaluación de las advertencias sanitarias y los sellos fiscales en Polonia y ayudan a los investigadores a determinar la legalidad de las cajetillas de cigarrillos (Gráfico 1).

La mayoría de las características de la cajetilla son fácilmente identificables de forma visual. Las características menos obvias de la cajetilla requieren atención o equipo adicional de expertos. Los fabricantes de cigarrillos ilícitos pueden colocar en sus cajetillas sellos fiscales falsos o extranjeros. Un experto perito podría determinar si el sello fiscal es auténtico. De forma alternativa, los investigadores pueden devolver las cajetillas a las autoridades de impresión de sellos, que pueden evaluar la autenticidad del sello fiscal. Con un equipo especializado, algunas cajetillas falsificadas se pueden distinguir mediante un escáner de microscopía de luz e irradiación ultravioleta que puede identificar las características distintivas. Estas características incluyen el tipo de tinta utilizada o el método de producción de la cajetilla (Kurti *et ál.*, 2017). Más adelante en el capítulo de la Sección 3.4.4 se ofrece más información sobre el examen pericial de la cajetilla.

El Gráfico 2 muestra un ejemplo del proceso de examen de una cajetilla. La lista de características de la cajetilla para el examen que se presenta en este gráfico no es exhaustiva y es probable que difiera según el país. Además de las advertencias sanitarias, otras marcas que se requieren en las cajetillas pueden incluir un número telefónico de ayuda para dejar de fumar, un número de identificación único para el seguimiento y localización de la cajetilla, así como el precio, el contenido de alquitrán y la información sobre el mercado de destino. Los investigadores pueden encontrarse con varios otros problemas durante el proceso de identificación de la legalidad de la cajetilla. A continuación, se exponen algunos ejemplos.

En los Estados Unidos (EE. UU.), la mayoría de los estados exigen sellos fiscales en las cajetillas de cigarrillos. Sin embargo, el sello fiscal se coloca en un empaque de celofán transparente en lugar de ir directamente en la cajetilla. Esto dificulta la determinación de si se han pagado los impuestos correspondientes, ya que la mayoría de los fumadores retiran el envoltorio de celofán inmediatamente después de abrir la cajetilla (Barker *et ál.*, 2016). Es muy poco probable que se encuentren cajetillas desechadas que también tengan un empaque de celofán intacto. Por consiguiente, las investigaciones sobre el comercio ilícito realizadas en los Estados Unidos requieren evaluaciones adicionales de las características de la cajetilla para determinar la legalidad de esta, además de los sellos fiscales.

Gráfico 2 Proceso del análisis de cajetillas



Brasil puso en marcha un sofisticado sistema de seguimiento y localización de sellos fiscales. Debería ser sencillo determinar si se pagó el impuesto brasileño adecuado por una cajetilla de cigarrillos determinada. Sin embargo, cuando se abre la cajetilla, el sello fiscal del impuesto al consumo, que contiene el número de identificación único de la cajetilla, se rompe. Lamentablemente, este diseño hace que el uso de las características de seguimiento y localización no sea adecuado para los investigadores que identifican las cajetillas que no cumplen con las normas, a menos que las cajetillas se obtengan antes de abrirlas.

Asimismo, las advertencias sanitarias de las cajetillas pueden no ser adecuadas para identificar la legalidad. Por ejemplo, la Directiva de Productos de Tabaco de la Unión Europea (UE) exige que las cajetillas de todos los Estados miembros lleven advertencias sanitarias tanto en forma de imágenes como de texto (Unión Europea, 2014). La UE proporciona la lista completa de posibles textos e imágenes (Comisión Europea, 2016). Los países que comparten el mismo idioma tienen cajetillas que son casi indistinguibles entre sí (por ejemplo, Alemania y Austria), ya que estas cajetillas muestran idénticas advertencias sanitarias y pictóricas. Los investigadores en esas circunstancias deben prestar especial atención a otros detalles de la cajetilla (por ejemplo, en el caso de Alemania y Austria, las cajetillas llevan diferentes mensajes para dejar de fumar) para distinguir entre ellas.

Los ejemplos anteriores sirven para ilustrar que las pruebas estadísticas no son intrínsecamente incorrectas, pero la aplicación de estas puede serlo. La investigación debe elegir la prueba estadística correcta, ya que éstas son simplemente herramientas. La sensibilidad (capacidad de detectar verdaderos resultados positivos) y la especificidad (capacidad de detectar verdaderos resultados negativos) son conceptos importantes en la caja de herramientas del estadista. Estos conceptos miden el rendimiento de la clasificación binaria. En una prueba para detectar cajetillas ilegales, la especificidad se refiere a la capacidad de la prueba para detectar si una cajetilla es legal, y la sensibilidad se refiere a la capacidad de la prueba para detectar si la cajetilla es ilegal.

Intuitivamente, un sello fiscal podría ser un buen indicador de la legalidad de la cajetilla. Sin embargo, como se mencionó anteriormente con respecto a los sellos fiscales en EE. UU., los sellos fiscales pueden faltar en las cajetillas legales (véase también Stoklosa y Ross, 2014). Durante el proceso de recolección de las cajetillas desechadas, los investigadores pueden encontrar sellos faltantes que fueron arrancados durante la apertura de la cajetilla o deslavados por la lluvia. Una prueba que clasifique cada cajetilla sin sello fiscal como una cajetilla ilegal tendría una baja especificidad. La baja especificidad se debe a que se especifican erróneamente todas las cajetillas como ilegales si no muestran un sello fiscal, lo que no es una forma precisa de medir la legalidad de la cajetilla.

Además, el bajo costo de los cigarrillos ilícitos podría utilizarse como indicador de legalidad. Es probable que las cajetillas ilegales sean significativamente más baratas que sus homólogos de fabricación legal. Sin embargo, una prueba que clasifique solo las cajetillas baratas como ilegales, sin tener en cuenta otras características de la cajetilla, podría pasar por alto las cajetillas ilegales que se venden a un precio superior. Esta prueba tiene una sensibilidad baja, ya que la proporción de cajetillas ilícitas correctamente identificadas es baja.

La aplicación de cada prueba implica compensaciones. Lo ideal sería que los investigadores pudieran seleccionar una prueba que ofrezca una alta sensibilidad y especificidad. Sin embargo, en la práctica, esto no suele ser así. Si los investigadores desean estimar el comercio ilícito de cigarrillos como el peor de los casos y están dispuestos a aceptar la posibilidad de que algunas cajetillas legales estén mal especificadas, entonces deberían utilizar una prueba de baja especificidad y alta sensibilidad. Este ejemplo destaca la necesidad de recopilar datos precisos, ya que las pruebas solo pueden ser tan buenas como los datos en los que se están comprobando.

Cabe señalar también que el incumplimiento de la cajetilla no siempre equivale a la elusión o evasión de impuestos. Por ejemplo, si se ha introducido una nueva advertencia sanitaria y ahora se exige en todas las cajetillas del país, pero siguen circulando algunas cajetillas antiguas con las versiones anteriores de la advertencia, esas cajetillas antiguas pueden estar sujetas al pago de impuestos y, al mismo tiempo, no cumplir con los requisitos de advertencia sanitaria. En tal caso, los resultados del cumplimiento de las advertencias sanitarias deben evaluarse e informarse por separado del cumplimiento de las leyes fiscales.

Al seleccionar las características de la cajetilla para su examen, también es importante tener en cuenta que la recopilación de datos requiere mucho tiempo y es costosa. Las mejores preguntas de la encuesta apuntan concisamente a respuestas que ayudan a responder la(s) pregunta(s) de la investigación. Preguntas de encuesta bien elaboradas pueden responder a la(s) pregunta(s) de la investigación con la mínima cantidad de información disponible. Puede resultar tentador reunir tanta información de las cajetillas de cigarrillos como sea posible, pero no toda la información de las cajetillas puede ser pertinente para responder a la(s) pregunta(s) de la investigación. La recolección de datos irrelevantes puede desperdiciar recursos valiosos sin añadir valor a la investigación.

Un ejemplo ilustrativo de este principio procede de México, donde la Ley General de Control del Tabaco en México define una lista de diez características de la cajetilla de cigarrillos que toda cajetilla de cigarrillos legal debe cumplir (Sáenz de Miera Juárez *et ál.*, 2020). Sin embargo, cuando los investigadores comenzaron un estudio sobre el comercio ilícito, se dieron cuenta de que llevaría un tiempo prohibitivo reunir información sobre las diez características de la cajetilla. Además, el aumento de las características de las cajetillas recolectadas también aumenta la probabilidad de que se produzcan errores de especificación debido a los daños en las cajetillas. Los investigadores recopilaron información sobre todas las características de las cajetillas para una submuestra (aproximadamente el 10 % de las cajetillas recopiladas). Basándose en la información de esa submuestra, los investigadores determinaron que tres características eran suficientes para distinguir con gran precisión las cajetillas de cigarrillos legales e ilegales. Por lo tanto, para el resto de la muestra (aproximadamente el 90%), se recopilaron datos sobre solo tres características de la cajetilla (Sáenz de Miera Juárez *et ál.*, 2020). De esta manera, los investigadores redujeron significativamente los costos de la investigación al disminuir la cantidad de datos que debían recogerse de cada cajetilla de cigarrillos sin sacrificar la efectividad.

3.2.3 Elección del método de investigación

Los investigadores tienen a su disposición dos métodos principales. Según las circunstancias específicas, los investigadores tal vez deseen emplear el método de análisis de cajetillas de cigarrillos del fumador o el método de recolección de cajetillas desechadas o utilizar ambos simultáneamente. El método de análisis de cajetillas de cigarrillos del fumador consiste en recopilar datos de las cajetillas de productos de tabaco durante una encuesta entre los consumidores de tabaco. En la encuesta también se les puede preguntar a los fumadores sobre su producto, como el precio del producto y el lugar de compra. Durante la recolección de cajetillas desechadas, los datos se recopilan de cajetillas que previamente se han desechado al suelo en lugares públicos. Los datos recopilados de la cajetilla se pueden aumentar con información sobre las localidades donde se encontró la cajetilla.

Es necesario comprender las ventajas y desventajas de cada método para aplicar adecuadamente las metodologías a la pregunta de investigación, las cuales se examinan a detalle a continuación. También se

examinan brevemente otros métodos de investigación. Se trata de la encuesta de cajetillas de los vendedores, en la que las cajetillas se obtienen de los vendedores al por menor, así como de la recolección de cajetillas debidamente desechadas de las oficinas, los hogares y los contenedores de basura públicos. Los cuatro métodos de investigación están definidos y explicados detalladamente en Ross (2015).

3.2.3.1 Ventajas de la encuesta de cajetillas de fumadores

El uso del análisis de cajetillas de fumadores tiene la ventaja de la amplitud de información disponible durante la fase de recolección. Como este método requiere una interacción directa con el fumador, las características del fumador se recopilan al mismo tiempo que se obtiene la información de la cajetilla. Esta característica de contacto directo está ausente en el método de recolección de cajetillas desechadas. Los investigadores pueden recopilar información sobre las características socioeconómicas del fumador, sus patrones de consumo de tabaco y el precio y el lugar de compra de la cajetilla. La información adicional puede ayudar a los investigadores a distinguir entre la elusión y la evasión de los impuestos al tabaco. La observación directa de los fumadores también ayuda a los investigadores a determinar quiénes son los fumadores de cigarrillos ilícitos y dónde se compran estos tipos de cigarrillos.

Una ventaja clara del método de encuesta de análisis de cajetillas de cigarrillos de fumadores es que los investigadores pueden determinar con mayor facilidad y precisión la legalidad de la cajetilla. Es frecuente que los turistas y los trabajadores extranjeros traigan consigo cigarrillos para su consumo personal. Estos cigarrillos importados son legales siempre y cuando su cantidad se mantenga dentro de los límites legales. El método de recolección de cajetillas desechadas no puede distinguir directamente entre las cajetillas importadas legal o ilegalmente por los turistas. Dado que la encuesta sobre el análisis de cajetillas de cigarrillos implica una interacción directa con los usuarios del producto, los investigadores pueden determinar el origen y el precio de la cajetilla. Disponer de esta información ayuda a los investigadores a distinguir entre la elusión fiscal legal y la evasión fiscal ilegal.

El método de recolección de cajetillas desechadas puede estar potencialmente sesgado, ya que su método depende de examinar las cajetillas que se tiraron al suelo en lugar de ser desechadas adecuadamente. Los resultados del método de recolección de cajetillas desechadas están sesgados cuando los fumadores de cajetillas legales e ilícitas las desechan con probabilidades significativamente diferentes. Por ejemplo, es posible que los fumadores de algunas localidades no se sientan cómodos fumando cajetillas ilícitas en público. Aunque los estudios que utilizaron tanto el examen de las cajetillas de fumadores como la recolección de cajetillas desechadas no encontraron diferencias significativas en la tasa de desecho (Merriman, 2010; Stoklosa y Ross, 2014; Sáenz de Miera Juárez *et ál.*, 2020), existe la posibilidad de que las cajetillas ilegales se desechen de manera no aleatoria. Por otra parte, el método de estudio del análisis de cajetillas de cigarrillos de fumadores permite investigar a todas las cajetillas.

También se deben considerar las normas culturales asociadas con el tirar basura en la calle. Si los fumadores con determinadas características son más propensos a tirar la basura en la calle, la recolección de cajetillas desechadas podría no representar las cajetillas que fuma la población en general. Esto hace que sea preferible el método de encuesta de análisis de cajetillas de fumadores, ya que permite controlar el análisis para determinar las características de los fumadores, como la edad, el sexo y la situación socioeconómica (por ejemplo, mediante un muestreo estratificado como se explica más adelante en este capítulo). El control de esos factores con el método de encuesta de análisis de cajetillas de fumadores puede eliminar los sesgos de las características de los fumadores que no se pueden eliminar de la recolección de cajetillas desechadas.

Por último, si bien es difícil realizar una encuesta sobre las cajetillas desechadas que abarque la superficie de todo un país, especialmente las zonas rurales, sigue siendo factible realizar una encuesta representativa a nivel nacional de los fumadores. Un ejemplo destacado de una encuesta representativa a nivel nacional que incorpora el análisis de cajetillas de cigarrillos es la Encuesta Mundial de Tabaquismo en Adultos (por ejemplo, en Polonia y Ucrania) (Centros para el Control y la Prevención de Enfermedades de los Estados Unidos, CDC, por sus siglas en inglés, 2019).

3.2.3.2 Ventajas del método de recolección de cajetillas desechadas

El método de recolección de cajetillas desechadas genera una estimación objetiva del comercio ilícito, la cual es una de sus principales ventajas. A diferencia de las encuestas de análisis de cajetillas de fumadores, las recolecciones de cajetillas desechadas no presentan problemas asociados con la autoevaluación. Los fumadores pueden no estar dispuestos a revelar sus hábitos de fumar debido a normas sociales o culturales. Los fumadores también pueden no estar dispuestos a admitir que fuman cigarrillos ilícitos, o pueden no querer mostrar su cajetilla de cigarrillos. De esta manera, el método de recolección de cajetillas desechadas tiene el potencial de reportar información más precisa sobre los hábitos tabáquicos.

El método de análisis de cajetillas desechadas también es fácil de aplicar. Incluso en las ciudades limpias, los recolectores de cajetillas no suelen tener problemas para recoger la cantidad deseada de cajetillas desechadas. Los estudios de recolección de cajetillas desechadas no involucran a sujetos humanos, por lo que no se requiere una revisión prolongada por parte de un comité de ética. Si el tiempo es esencial, como cuando los investigadores quieren medir el comercio ilícito antes y después de una intervención de control del tabaco o de un aumento del impuesto al consumo de tabaco, el método de recolección de cajetillas desechadas permite una mayor flexibilidad en la aplicación del estudio. Incluso en el caso (muy improbable) de que se requiera un examen ético, este debe ser sencillo. Algunos países pueden exigir un examen ético con el fin de evaluar los riesgos a los que podrían estar expuestos los investigadores.

Otra ventaja de los análisis de cajetillas desechadas es que los investigadores pueden pasar tanto tiempo en una cajetilla desechada como sea necesario para recoger los datos deseados. Las cajetillas desechadas se pueden recolectar y enviar a un centro de investigación con el equipo y los recursos adecuados para evaluar la legalidad de las cajetillas. Por ejemplo, una investigación pericial de expertos para determinar las cajetillas falsificadas (véase la Sección 3.4.4 para una descripción de cómo realizar tal prueba) sería imposible sin una cajetilla física. Además, las cajetillas recolectadas pueden almacenarse y revisarse posteriormente. Esto es particularmente útil porque las mismas cajetillas se pueden reutilizar para un estudio posterior que pueda examinar diferentes características de la cajetilla. La capacidad de almacenar las cajetillas recolectadas también puede aumentar la transparencia del estudio, ya que las cajetillas se pueden volver a examinar cuando se solicite la prueba de los resultados. Aunque es posible comprarles o solicitarles las cajetillas a los fumadores en una inspección de cajetillas de fumadores, no todos los fumadores estarían de acuerdo en vender o dárselas a los encuestadores. Además, aunque sea posible tomar una fotografía de la cajetilla durante la encuesta de fumadores, estas fotografías pueden no ser suficientes si se necesitan pruebas adicionales.

Además de ser más sensibles al tiempo, los estudios de recolección de cajetillas desechadas también son menos costosos que las inspecciones de cajetillas de las encuestas a fumadores. En la encuesta de análisis de cajetillas se requieren encuestadores expertos. Entre las habilidades que se requieren figuran la capacidad de acercarse a una persona y establecer una relación de comunicación, leer y hablar claramente en el idioma nativo, y escuchar atentamente y manteniéndose neutral durante toda la entrevista. Estos encuestadores necesitan capacitación en el uso de la guía del encuestado, los materiales visuales utilizados en la entrevista y, en el caso de las entrevistas personales asistidas por computadora (CAPI, por sus siglas en inglés) u otras herramientas, necesitan capacitación para utilizar el software de la encuesta. Las empresas de investigación de mercados que cuentan con equipos de encuestadores capacitados pueden cobrar precios altos.

El método de recolección de cajetillas desechadas, por otro lado, no requiere el mismo nivel de habilidad. Las habilidades que se requieren de los recolectores de cajetillas se limitan a la capacidad de caminar por una ruta designada, recolectar las cajetillas desechadas y registrar la ubicación de la recolección, o clasificar la basura y registrar su origen. Por lo tanto, los servicios de las empresas de investigación de mercados a menudo no son necesarios para el método de recolección de cajetillas desechadas. Los encuestadores para este tipo de estudio se pueden reclutar entre los trabajadores menos calificados a un costo menor que el asociado a encuestas a fumadores. Por ejemplo, en algunos estudios anteriores, se reclutaron estudiantes universitarios para que recolectaran cajetillas desechadas.

3.2.3.3 Otros métodos de análisis de cajetillas

Hay otros métodos de análisis de cajetillas disponibles para los investigadores que proporcionan una validación cruzada útil de los resultados de la encuesta a fumadores y de recolección de cajetillas desechadas y revelan más información sobre el comportamiento de los fumadores. El primer método alternativo se conoce como la encuesta al vendedor de cajetillas. En muchos países, un gran porcentaje de los cigarrillos se venden en forma de cigarrillos individuales. Por lo tanto, los fumadores de cigarrillos individuales no tienen una cajetilla que los investigadores puedan examinar. Si los vendedores de cigarrillos individuales desechan sus cajetillas de cigarrillos al mismo ritmo que los fumadores que compran cajetillas, entonces un análisis de cajetillas desechadas es adecuado para estimar la prevalencia de los cigarrillos ilícitos en esos países. Sin embargo, no siempre es así. Dado que los puestos se instalan por la mañana, y en algunos países los vendedores permanecen en el mismo lugar durante todo el día, los vendedores suelen disponer de un sistema adecuado de desecho de sus cajetillas usadas. En esos casos, el método de recolección de cajetillas desechadas pasaría por alto las cajetillas debidamente desechadas por los vendedores de cigarrillos individuales.

En la encuesta a vendedores de cajetillas, el vendedor de cigarrillos sueltos aparta las cajetillas de cigarrillos vacías para que los encuestadores las recojan al final del día. La encuesta a vendedores de cajetillas de cigarrillos se implementó por primera vez en la India, donde el 55% de los cigarrillos se venden sueltos (John y Ross, 2018). En este estudio, a cambio de una pequeña recompensa monetaria, los vendedores de cigarrillos individuales a lo largo de rutas predeterminadas en ocho ciudades de la India depositaron cajetillas de cigarrillos vacías en bolsas provistas, que se recogieron al final del día.

Aunque la participación directa de los consumidores no forma parte de este tipo de estudio, los encuestadores pueden reunir información de los vendedores sobre las marcas, las cantidades y los precios de los cigarrillos. Sin embargo, es importante recordar que la información que se recoge de los vendedores puede estar sujeta a prejuicios relacionados con la disposición del vendedor a responder y la veracidad de esas respuestas. Si la venta de cigarrillos ilegales conlleva una pena o es un delito, el vendedor podría ser especialmente reacio a darle sus cajetillas vacías al encuestador o a responder a sus preguntas. Por esta razón, la información proporcionada por los vendedores sobre las marcas de cigarrillos, precio y cantidad es secundaria y sirve solo como complemento o comprobación de la solidez del objetivo principal de recoger las cajetillas de cigarrillos de los vendedores. La información proporcionada por el vendedor no se debe utilizar para informar sobre la prevalencia de cajetillas ilícitas antes de su examen, porque la proporción de cajetillas ilícitas entre las cajetillas desechadas por el vendedor proporciona la información más precisa. Por último, como el vendedor podría ser reacio a proporcionar las cajetillas desechadas a los encuestadores, los investigadores deberían considerar la posibilidad de incentivar la participación en la encuesta. Es importante que la recompensa monetaria sea pequeña y uniforme. En otras palabras, la recompensa monetaria no puede basarse en el número de cajetillas recolectadas, ya que puede incentivar a los vendedores a añadir cajetillas adicionales. Esto sesgaría los resultados de manera sustancial y no aleatoria.

Aunque el método de la encuesta a vendedores de cajetillas no es el centro de este conjunto de herramientas, varios aspectos del diseño del estudio que se describen más adelante en este capítulo se aplican también directamente a este método. Debido a la similitud entre este método y el de recolección de cajetillas desechadas, la planeación del estudio, la ejecución y el análisis de los datos están directamente relacionados con la encuesta de cajetillas desechadas de los proveedores. Por ejemplo, en ambos métodos, los encuestadores deben seguir rutas predeterminadas, que se seleccionan en función del lugar en que los fumadores residen y compran sus cigarrillos. La principal diferencia es que en el método de recolección de cajetillas desechadas todas las cajetillas desechadas se recogen a lo largo de la ruta, mientras que en la encuesta a vendedores de cajetillas, las cajetillas se recogen de los vendedores a lo largo de la ruta.

La combinación de la recolección de cajetillas desechadas y los métodos de encuesta a los vendedores proporciona una visión más completa del comportamiento general del fumador. Los consumidores de cigarrillos individuales se diferencian de los consumidores que compran en cajetillas, ya sea por la frecuencia del consumo o por el nivel de ingresos. Si el 30% de los cigarrillos que se fuman proceden de vendedores de cigarrillos individuales mientras que el 70% se compran en cajetillas, entonces la prevalencia del comercio ilícito puede estimarse calculando el promedio ponderado de la prevalencia de cigarrillos ilícitos de ambas fuentes.

El último método que se examina en este capítulo es el de la recolección de cajetillas debidamente desechadas. Es posible que las cajetillas desechadas adecuadamente capten a un consumidor diferente del método de recolección de cajetillas desechadas y, por lo tanto, haga que los dos métodos sean complementarios. La diferenciación entre ambos puede proporcionar alguna información sobre los tipos de fumadores. El método de recolección de cajetillas desechadas recoge las cajetillas en el exterior, después de que el fumador haya terminado su cajetilla. Esto significa que la recolección de cajetillas desechadas no incluye a las personas que predominantemente fuman, y se deshacen de sus cajetillas, dentro de sus casas, oficinas u otros establecimientos.

Para hacer frente a esta omisión, los investigadores se pueden centrar en la recolección de las cajetillas desechadas adecuadamente que se encuentran en los contenedores y receptáculos de basura situados cerca de los lugares de trabajo y las residencias. El método se ha utilizado en varios estudios, entre ellos los realizados en Chicago (Merriman, 2010), en la ciudad de Nueva York (Consroe *et ál.*, 2016) y en París (Lakhdar, 2008).

Un inconveniente del método de recolección de cajetillas debidamente desechadas es la dificultad de generar una muestra representativa. Para que la muestra sea representativa, es necesario recolectar basura de una amplia gama de lugares. Aunque teóricamente es fácil, en la práctica es difícil de lograr. Las empresas de recolección de basura se diferencian por los lugares públicos y las residencias, y varias empresas diferentes pueden operar en la misma zona abarcando diferentes sectores. Esto hace que sea difícil obtener una muestra representativa, o una muestra sesgada para la cual los investigadores puedan explicar el sesgo. Es necesario colaborar con los grupos de recolección de basura, ya que generalmente tienen acceso exclusivo a los contenedores de basura. Cuando trabajan con estas empresas, los investigadores interceptan la basura en la estación de transferencia antes de que la basura vaya al vertedero. Las grandes empresas de recolección de basura recogen grandes cantidades de basura, lo que complica aún más las cosas, ya que se desconoce el origen exacto de la basura.

Elegir el método de recolección de desechos adecuadamente significa tener que clasificar la basura recolectada, ya que por lo general no se preclasifica. El almacenamiento adecuado de la basura, como refrigeradores y congeladores también sería necesario, ya que la clasificación a través de las bolsas de basura consume mucho tiempo. Como las bolsas de basura también contienen productos perecederos, el material orgánico comenzará a descomponerse bastante rápido. Minimizar los olores de la basura es necesario para la comodidad de los encuestadores. Es necesario tomar varias otras precauciones, ya que los encuestadores pueden encontrarse con materiales potencialmente peligrosos o nocivos durante el proceso de clasificación, como agujas médicas y objetos punzantes.

El método de recolección de cajetillas debidamente desechadas es el enfoque de análisis de cajetillas más utilizado, y es cuestionable si este método es más preciso que otros. El principal beneficio de utilizar este método es establecer si las características de los que las desechan son similares a las de los que no lo hacen. El alto costo y el valor desconocido que se obtiene de este método no lo hacen viable, por lo que no se examina más a fondo en este conjunto de herramientas, pero algunas partes de este conjunto de herramientas también son aplicables al método de recolección de cajetillas debidamente desechadas. Entre ellas figuran las secciones sobre las características de las cajetillas, el diseño de una muestra de zonas geográficas y la recopilación de datos de las cajetillas.

Los investigadores que realicen estudios utilizando estos métodos deben tener en cuenta sus fortalezas y debilidades. Como ya se ha mencionado, de ser posible, los investigadores deberían utilizar el otro o los otros métodos para validar de forma cruzada sus hallazgos para lograr el enfoque más sólido y preciso. La validación cruzada es útil porque cada método capta aspectos del comercio ilícito que otros métodos no pueden captar. Por ejemplo, el proceso de recolección de datos de las encuestas sobre cajetillas a fumadores es más caro que el método de recolección de cajetillas desechadas. Sin embargo, si se utilizan en conjunto, el método de recolección de cajetillas desechadas se puede utilizar en una amplia zona geográfica, mientras que la encuesta de análisis de cajetillas de fumadores se puede llevar a cabo en unas pocas subsecciones. De esta manera, se pueden utilizar los beneficios de ambos métodos y se mitigan los inconvenientes de cualquiera de ellos mediante la validación cruzada.

Este enfoque mixto fue utilizado recientemente por investigadores en Brasil. Se recolectaron cajetillas desechadas en cinco ciudades (Belo Horizonte, Campo Grande, Joao Pessoa, Río de Janeiro y Sao Paulo), y la encuesta de análisis de cajetillas de fumadores se llevó a cabo en dos de esas ciudades (Río de Janeiro y Sao Paulo). Además, los investigadores realizaron una encuesta sobre cajetillas desechadas adecuadamente de las residencias en Río de Janeiro. En combinación, los investigadores pudieron estimar resultados sólidos (Carvalho Figueiredo et ál., próxima publicación).

3.3 Muestreo

El diseño de la muestra es un aspecto importante del diseño del estudio y requiere algunos comentarios adicionales. Es probable que los investigadores con experiencia previa en encuestas se sientan cómodos con la ejecución de encuestas y el análisis de datos, pero los métodos de muestreo son relativamente diferentes entre los dos métodos principales presentados anteriormente, y esto puede requerir una aclaración.

Una muestra se define como un subconjunto de una población. Por lo tanto, antes de considerar una muestra, los investigadores deben definir claramente la población de estudio, la cual se examina a continuación.

3.3.1 Población de estudio

Los investigadores deben definir explícitamente la población de estudio y luego considerar cómo encaja la muestra en ella. Cuando la población de estudio está mal definida, es improbable que la muestra sea representativa de la verdadera población (Sudman, 1983). La definición de la población de estudio en los estudios sobre el comercio ilícito requiere especialmente una planeación cuidadosa y es esencial para llegar a estimaciones precisas. Esta sección incluye puntos importantes que se deben considerar al definir la población de estudio en los estudios de comercio ilícito.

En primer lugar, un hecho importante que hay que recordar al definir la población de estudio es que los adultos tienen diferentes prevalencias de tabaquismo. En la mayoría de los lugares, los adultos con mayores ingresos tienden a tener menor prevalencia de tabaquismo (Drope et ál., 2018). Por lo tanto, una muestra representativa tiene en cuenta la distribución de los fumadores adultos en lugar de distribuir a los sujetos de las encuestas de manera uniforme. Si las encuestas se realizan de manera uniforme, se creará una representación excesiva de las personas en los niveles socioeconómicos más altos, que históricamente tienen prevalencia de tabaquismo más baja. Los investigadores que diseñen muestras de estudios sobre el tabaco ilícito deben centrarse en el fumador típico y en el lugar donde viven y trabajan estos fumadores.

Otro punto importante que hay que considerar es que el objetivo final del estudio es reunir información sobre los cigarrillos. Los fumadores son la fuente de esta información. Se hace una encuesta entre los fumadores para obtener información sobre sus hábitos tabáquicos a fin de estimar el comercio ilícito de tabaco. Los cigarrillos fumados constituyen la población de estudio, no los fumadores. Esta puede ser una distinción importante cuando el comportamiento tabáquico difiere entre los fumadores de cigarrillos ilícitos y los fumadores habituales. Es posible, por ejemplo, que los fumadores de cigarrillos ilícitos fumen más cigarrillos al día que el fumador promedio. Las estimaciones de la prevalencia de los cigarrillos ilícitos deben tener en cuenta estas diferencias de comportamiento de los fumadores.

Además, al realizar investigaciones sobre el comercio ilícito de tabaco, los investigadores deben tener en cuenta que la población de estudio son todos los cigarrillos que se compran y consumen, no solo los cigarrillos ilícitos. Con frecuencia, los investigadores del comercio ilícito tienden a centrarse exclusivamente en los cigarrillos ilícitos y a menudo se centran incorrectamente en la recopilación de datos en zonas conocidas por ser zonas de alto tráfico de cigarrillos ilícitos. Si bien el muestreo de los puntos de conflicto puede formar parte del diseño de un estudio, la recopilación de datos de esta manera sobrestimaría la proporción de cigarrillos ilícitos. La publicación de investigaciones realizadas de esta manera inflará incorrectamente las estimaciones del consumo ilícito de cigarrillos.

El siguiente punto para considerar es que los estudios sobre el comercio ilícito se deben centrar en el consumo actual de cigarrillos. El comportamiento pasado no es necesariamente informativo sobre el comportamiento actual. Lo ideal sería que la encuesta se centrara en la última cajetilla de cigarrillos comprada o consumida. Las preguntas sobre el pasado no se deben usar a menos que se recopile información sobre el año y el mes de ese comportamiento. La ambigüedad introduce un sesgo si estas preguntas no son específicas en cuanto al tiempo. Las preguntas ambiguas recogerán datos inexactos. Cuando se pregunte a los fumadores sobre su comportamiento como fumadores en el pasado, las preguntas se deben hacer con mucho cuidado. Por ejemplo, la encuesta del Eurobarómetro pregunta a los fumadores si les han ofrecido cigarrillos del mercado negro. Las posibles respuestas son las siguientes: rara vez (menos de una vez al mes); ocasionalmente (de una a tres veces al mes); regularmente (una vez a la semana o más frecuentemente); o nunca (Comisión Europea, 2016). Para responder a esta pregunta con sinceridad, una persona a la que se le ofreció una cajetilla de cigarrillos ilícitos hace una década no puede elegir la opción de “nunca” en el cuestionario. Por lo tanto, la encuesta del Eurobarómetro no está bien diseñada. Asimismo, en algunas encuestas se pregunta si los fumadores han comprado alguna vez cigarrillos ilícitos. Eso tampoco es preciso y esas preguntas se deben evitar.

La reducción de la población de estudio apropiada también requiere una consideración cuidadosa. Como ya se ha mencionado, las estimaciones del comercio ilícito de tabaco se deben centrar en todos los cigarrillos que se compran y consumen. Sin embargo, los estudios sobre la disponibilidad de cigarrillos ilícitos se centran en una población diferente a la de los estudios que estiman el nivel de comercio ilícito de cigarrillos. Una población de estudio incorrectamente definida conducirá a resultados sesgados. En la mencionada encuesta del Eurobarómetro, el objetivo del estudio era medir la penetración del mercado negro (Comisión Europea, 2016). La investigación se centró en la medición de la prevalencia del ofrecimiento de cigarrillos ilícitos. La encuesta del Eurobarómetro reveló que Luxemburgo (el 94% de los encuestados declaró que nunca se le habían ofrecido cigarrillos ilícitos) ocupaba el lugar más alto en cuanto al comercio legal de tabaco, y Lituania y Bulgaria (ambos el 64%) el lugar más bajo. El informe concluye que “geográficamente, los cigarrillos del mercado negro tienen una mayor penetración en los países de Europa del Este que en otros lugares”.

Es probable que el estudio haya dado respuestas sesgadas, impulsadas por parámetros mal definidos de la población estudiada. La probabilidad de que se le ofrezca un cigarrillo ilícito suele ser diferente en los mercados locales (bazares) que en las cadenas de supermercados. Además, el comportamiento de compra de alimentos puede ser marcadamente diferente entre países y por lo tanto, puede ser un factor de confusión. Es probable que la prevalencia de las compras en cadenas de supermercados en Luxemburgo sea mayor, mientras que probablemente la prevalencia de las compras en los mercados locales de Lituania y Bulgaria sea mayor. Cuando la probabilidad de que se ofrezcan cigarrillos ilícitos es mayor en los mercados locales, el estudio será sesgado si los datos no están debidamente ponderados para tener en cuenta esta diferencia. En consecuencia, los datos obtenidos no confirman necesariamente que la penetración del mercado negro sea mayor en los países de Europa del Este.

3.3.2 Geografía

La siguiente etapa de planeación de la investigación es la selección del área de estudio a cubrir. El objetivo del estudio de cada cajetilla de cigarrillos es producir una muestra representativa de cajetillas de cigarrillos fumados dentro de una unidad administrativa más grande (por ejemplo, una ciudad, una provincia o un país). El financiamiento de los estudios de investigación independientes suele ser más limitado. Las limitaciones presupuestarias más estrictas a menudo hacen inviable la toma de muestras en todo el país. Las encuestas de análisis de cajetillas rara vez son representativas a nivel nacional (entre las raras excepciones se encuentra Fix *et ál.*, 2014), o tienen una cobertura nacional (por ejemplo, Barker *et ál.*, 2016; Little *et ál.*, 2019).

Es difícil llevar a cabo una recolección de cajetillas desechadas representativa a nivel nacional. Para que este método sea representativo a nivel nacional, los investigadores tendrían que recoger cajetillas de varias zonas diferentes del país, basándose en un marco de muestreo cuidadosamente diseñado que asegurara la representatividad de la muestra. Esto incluiría recolecciones en las ciudades donde la población es más densa, así como en las zonas rurales. La probabilidad de encontrar cajetillas desechadas en las zonas menos pobladas disminuye considerablemente a medida que se distribuyen menos personas en una zona más extensa. Para cubrir

adecuadamente las zonas rurales, los costos de los estudios aumentarían drásticamente, ya que se requieren más horas de trabajo para cubrir estas zonas.

En lugar de aspirar a una muestra representativa a nivel nacional, los estudios de recolección de cajetillas desechadas suelen limitarse a un número de ciudades. Aunque esto limita el alcance del estudio, el estudio sigue siendo válido siempre y cuando se recojan muestras representativas de las ciudades seleccionadas. Entre los requisitos adicionales para los estudios de recolección de cajetillas desechadas figuran que las ciudades deben estar repartidas por todo el país; debe haber un número suficiente de personas que vivan en cada ciudad; y los consumidores de tabaco de cada ciudad deben reflejar a los consumidores promedio del país en su conjunto. Las estimaciones de ese estudio ofrecen una visión imparcial del comercio ilícito de tabaco en el país, especialmente si la población urbana estudiada representa una gran parte de la población total.

Por ejemplo, en un estudio reciente de Colombia se estima el nivel de elusión y evasión de impuestos al tabaco. Los investigadores utilizaron el método de análisis de cajetillas de cigarrillos en varias ciudades para obtener una muestra representativa a nivel nacional (Maldonado *et ál.*, 2018). La encuesta se realizó en cinco ciudades colombianas, que representan el 63% del mercado de cigarrillos del país. La muestra se diseñó inteligentemente para que coincidiera con la distribución por edad y sexo de los fumadores de una encuesta representativa nacional diferente. Otro ejemplo es un estudio mexicano que utilizó un enfoque similar (Sáenz de Miera Juárez *et ál.*, 2020). Tanto la encuesta sobre el análisis de cajetillas de cigarrillos como la recolección de cajetillas desechadas se llevaron a cabo en ocho ciudades del país. Asimismo, otra encuesta sobre las cajetillas de los vendedores en la India recogió cajetillas vacías de los vendedores de cigarrillos individuales en ocho ciudades (John y Ross, 2018). Tanto el estudio de México como el de la India abarcaban una gran parte de la población total y tenían una vasta extensión geográfica.

3.3.3 Edad y otras características de las personas

Otro factor para tener en cuenta es que los investigadores deben establecer una edad mínima de corte para los sujetos de la encuesta. No se trata de una preocupación para el método de recolección de cajetillas desechadas, porque implica la recolección de cajetillas desechadas por todos los fumadores (ya que es imposible discernir las características de los fumadores a partir de las cajetillas solamente). Normalmente se requiere el consentimiento de los padres para realizar encuestas a menores porque no es éticamente apropiado preguntar sobre el comportamiento de los menores en relación con el tabaquismo. Cuando se fija una edad legal para fumar (normalmente 18 años), la edad límite para la encuesta debe ser la misma.

El estado de residencia de un fumador también requiere un valor límite. Es posible que los investigadores quieran identificar a los turistas y trabajadores temporales de su muestra, ya sea para excluirlos o para medir un aspecto diferente de la investigación. Los turistas y los trabajadores temporales no forman parte de la población general del país. Una consideración similar debería darse a las personas que viven fuera de los límites definidos de la ciudad o la jurisdicción. Si un porcentaje sustancial de trabajadores se desplazan a la ciudad estudiada, entonces se deben establecer parámetros apropiados para este tipo de ciudad.

3.3.4 Tamaño de la muestra

El siguiente paso en el proceso de muestreo es establecer el tamaño de la muestra. A menudo el tipo de pregunta de estudio determinará el tamaño de la muestra requerida (Sudman, 1983). No existe una regla estricta sobre el tamaño de muestra apropiado para los estudios sobre el comercio ilícito de tabaco. Para un estudio general, las estimaciones no tienen que ser particularmente exactas ya que estos estudios informan de estadísticas descriptivas. Este tipo de estudio sirve como un estudio exploratorio y esencialmente proporciona un primer vistazo al problema. Para los estudios que proporcionan un análisis más profundo, incluida la inferencia estadística, se necesitan muestras más grandes y más cuidadosamente seleccionadas. Los estudios requieren muestras más grandes y representativas si pretenden comprobar si sus estimaciones son estadísticamente diferentes de las de la industria tabacalera. Análogamente, las estimaciones a nivel de las ciudades sobre la elusión y la evasión fiscal en el consumo de tabaco también requieren una muestra suficientemente amplia. Para medir el incumplimiento fiscal puede ser necesaria una muestra aún mayor para medir adecuadamente los niveles en las diferentes ciudades.

El desgaste de la muestra también se debe tener en cuenta al establecer el tamaño de esta. Las bajas tasas de respuesta en las encuestas sobre cajetillas de fumadores son una fuente de desgaste. Los fumadores pueden ser incapaces o sentirse incómodos al responder a preguntas sobre su hábito tabáquico o al presentar sus cajetillas para ser analizadas. Un estudio piloto preliminar a pequeña escala les da a los investigadores una idea general de la tasa de desgaste de la muestra. Basándose en esta información, los investigadores pueden tener una mejor idea del tamaño de la muestra requerida, así como de la cantidad de tiempo necesario para completar cada encuesta.

Un estudio piloto también es importante para el método de recolección de cajetillas desechadas. El estudio piloto les da a los investigadores una idea general de cuántas cajetillas se pueden recolectar en cada área. Dado que hay una gran variación en la densidad de población, la prevalencia de tabaquismo y el comportamiento de desecho de cajetillas, el número de cajetillas desechadas también puede tener una gran variación. Por ejemplo, en un estudio de recolección de cajetillas desechadas realizado en Chicago, los recolectores tuvieron que caminar aproximadamente 135 millas (217 km) (Merriman, 2010) para recoger la muestra requerida. Un estudio similar realizado en México solo requirió que los recolectores caminaran alrededor de 52 millas (83 km) por ciudad (Sáenz de Miera Juárez, 2020). Sorprendentemente, ambos estudios recogieron casi el mismo número de cajetillas por ciudad (alrededor de 1.000). Un estudio piloto a pequeña escala les proporcionó a los investigadores la información necesaria para establecer la distancia requerida para recolectar el número apropiado de cajetillas.

La siguiente fórmula estadística ayuda a estimar el tamaño mínimo de la muestra para un estudio, dado el nivel de precisión deseado:

$$\hat{n} = \frac{Z_{1-\alpha/2}^2 \cdot P(1-P) \cdot DEFF}{\delta^2}$$

donde n es el tamaño de la muestra, Z es el estadístico Z para un determinado nivel de confianza (por ejemplo, la puntuación Z crítica es $Z_{1-0.05/2} = 1,96$ para el nivel de confianza estándar del 95%), P es la prevalencia esperada de cajetillas ilícitas (por ejemplo, $P = 0,2$ para una prevalencia esperada del 20%), $DEFF$ es un efecto de diseño que tiene en cuenta el diseño de la muestra compleja (véase más adelante), y δ es la precisión supuesta. La precisión supuesta también se denomina margen de error (por ejemplo, $\delta = 0,05$ si el margen de error supuesto es del 5%). Esta n es el tamaño mínimo de muestra necesario para una inferencia precisa y confiable. La fórmula de n supone una población infinita (muy grande). Así pues, en el caso de poblaciones grandes, el tamaño de la muestra no depende del tamaño de la población de estudio. Una n más pequeña no tendrá la precisión necesaria para una prueba de inferencia precisa en el nivel de confianza estadístico dado. Aunque las muestras más grandes no perjudicarán el estudio, los costos de este aumentarán innecesariamente sin beneficio adicional.

Cabe señalar varios detalles sobre esta fórmula. En primer lugar, el tamaño de muestra deseado depende de la prevalencia de cigarrillos ilícitos (P). El estudio piloto les dará a los investigadores una estimación aproximada del tamaño de P . Matemáticamente, el valor de $P(1 - P)$ es mayor cuando $P = 0,5$. El tamaño de la muestra será mayor cuando $P = 0,5$ (la mitad de las cajetillas son ilícitas). En segundo lugar, el factor de ajuste $DEFF$ se expresa como la relación de la varianza para el diseño de la muestra dada dividida por la varianza de la muestra efectiva del mismo tamaño. La muestra más efectiva se obtiene mediante un muestreo aleatorio. Para las muestras aleatorias simples, el numerador de la proporción $DEFF$ es igual a su denominador. Por lo tanto, $DEFF$ es igual a uno. En el caso de estructuras de muestra más complicadas, como las descritas en las siguientes secciones de este capítulo, la varianza en el numerador de la relación $DEFF$ es diferente de la varianza en el denominador, por lo que el factor $DEFF$ no será igual a uno.

Una desventaja del factor $DEFF$ es que para calcularlo se requiere la varianza de la muestra. Sin embargo, la varianza de la muestra no se puede determinar hasta que se haya realizado el estudio. Esto es poco práctico para muchos proyectos de investigación. A menudo se utiliza una regla empírica al seleccionar el factor $DEFF$. En el caso del muestreo por conglomerados, que es un diseño de muestreo que se utiliza a menudo en los estudios de examen de cajetillas, se suele suponer un efecto de diseño mayor o igual a 2 cuando no se dispone de estimaciones $DEFF$ (Bostoen y Chalabi, 2006; Maas y Hox, 2005).

Los investigadores mexicanos que llevaron a cabo el estudio de recolección de cajetillas desechadas utilizaron la fórmula anterior para determinar el tamaño apropiado de la muestra. El estudio supuso: $Z = 1,96$ (el valor crítico de la puntuación Z para el nivel de confianza del 95%), $P = 0,02$ (basado en la expectativa de que el 2% de las cajetillas son ilícitas), $DEFF = 2$ (que permite la variabilidad en el número de cajetillas recolectadas por unidad de muestreo y la prevalencia de cajetillas ilícitas en cada unidad de muestreo), y $\delta = 0,015$ para el supuesto margen de error del 1,5%, lo que arroja un tamaño de muestra de 670 cajetillas por ciudad. El estudio piloto determinó que los recolectores deben encontrar unas 8 cajetillas por kilómetro. Así pues, la distancia mínima de viaje necesaria para recolectar la cantidad necesaria es de 52 millas (83 km) por ciudad (Sáenz de Miera Juárez, 2020). El objetivo del estudio en México era seleccionar un tamaño de muestra que representara con precisión el mercado ilícito de cigarrillos a nivel de la ciudad. Si los investigadores hubieran querido comparar los niveles de comercio ilícito de cigarrillos en los distintos distritos de la ciudad, también se habrían tenido que elegir tamaños de muestra apropiados a nivel de distrito. Cuando los investigadores quieran realizar estudios de subpoblación (por ejemplo, por grupo étnico o género), la fórmula deberá utilizarse para cada subpoblación. El tamaño de la muestra tendría que ser inflado para la prueba de la muestra estratificada.

También se puede tomar un tamaño de muestra de un estudio realizado previamente. Si bien el uso de la fórmula para determinar el tamaño de la muestra proporciona a los investigadores un número de muestra preciso, la adopción de un tamaño de muestra de un estudio anterior no lo hace. En los estudios de alcance nacional, el tamaño de la muestra suele ser superior a 2.000 (por ejemplo, Barker *et ál.*, 2016; Little *et ál.*, 2019) y a veces hasta 8.000 (Sáenz de Miera Juárez *et ál.*, 2020). Las estimaciones a nivel de las ciudades suelen tener cada una alrededor de 1.000 observaciones (por ejemplo, Merriman, 2010; Sáenz de Miera Juárez *et ál.*, 2020). Una vez más, los investigadores que deseen obtener estimaciones a nivel de subciudad o de submuestra deben aumentar el tamaño de la muestra.

3.3.5 Diseño de la muestra

El diseño de la muestra define el conjunto de reglas que especifican cómo seleccionar la muestra. La investigación de encuestas, como los estudios de análisis de cajetillas de fumadores, suele utilizar un diseño de muestreo probabilístico. El principio subyacente en el muestreo de probabilidad es que la selección de cada elemento de la población tiene una probabilidad no nula. Las muestras aleatorias simples son el tipo más básico de muestras. Cada elemento de la población tiene igual probabilidad de ser seleccionado en la muestra. La estratificación divide la muestra en subpoblaciones mutuamente excluyentes y exhaustivas (por ejemplo, hombre/mujer, niño/adulto, etc.) llamadas estratos. Se toman muestras de cada subgrupo. Por último, el muestreo por conglomerados divide a la población en conglomerados, de los cuales se selecciona una muestra de conglomerados. En la siguiente sección se presenta cada una de estas técnicas de muestreo en el contexto de los estudios de análisis de cajetillas.

3.3.5.1 Muestreo aleatorio simple

El muestreo aleatorio simple es el diseño más sencillo. Un atributo de este método es la baja variabilidad de la muestra. En la práctica, el muestreo aleatorio simple es difícil de aplicar en encuestas grandes de población. Una característica fundamental del muestreo aleatorio simple es que la probabilidad de ser seleccionada en la muestra es igual entre observaciones. Dado que las personas no están distribuidas uniformemente en todo el país y las cajetillas de cigarrillos no están distribuidas uniformemente entre las personas, este criterio es difícil de cumplir. Una muestra aleatoria es poco probable cuando la gente se acerca a la calle. La hora del día determina qué tipo de personas pueden ser encuestadas. Temprano en la mañana, la gente va en camino al trabajo. Más tarde, los investigadores pueden encontrar una mayor proporción de estudiantes universitarios que no están atados a un horario estricto de nueve a cinco. Este tipo de enfoque también puede extrañar por completo a las personas que rara vez salen de sus casas.

La aplicación adecuada del muestreo aleatorio puro requeriría que los investigadores poseyeran una lista completa de los elementos de la población (marco poblacional) de los que se extrae la muestra de forma aleatoria. Esta información les permitiría a los investigadores tomar muestras uniformes de los elementos de la población y, por lo tanto, satisfacer el criterio de probabilidad uniforme. Sin embargo, en el caso de las encuestas de análisis de cajetillas de fumadores, es muy poco probable que exista una lista completa de fumadores.

En el caso improbable de que exista una lista completa de hogares con al menos un fumador, puede que todavía no sea factible aplicar un muestreo aleatorio simple. Dado que cada hogar seleccionado tendría que ser visitado individualmente, la visita a todos los hogares seleccionados al azar podría ser logísticamente difícil o costosa, especialmente si están dispersos en grandes zonas geográficas. La tasa de desgaste también puede ser alta porque los residentes pueden no estar presentes en el momento de la visita. Además, las conductas de consumo de tabaco pueden variar entre los fumadores.

El muestreo aleatorio simple es fácil de aplicar en teoría, pero satisfacer adecuadamente el criterio es un desafío en la práctica. Los investigadores que realizan estudios de encuestas tienden a recurrir a otros métodos de selección de muestras: la estratificación y la conglomeración.

3.3.5.2 Muestreo estratificado

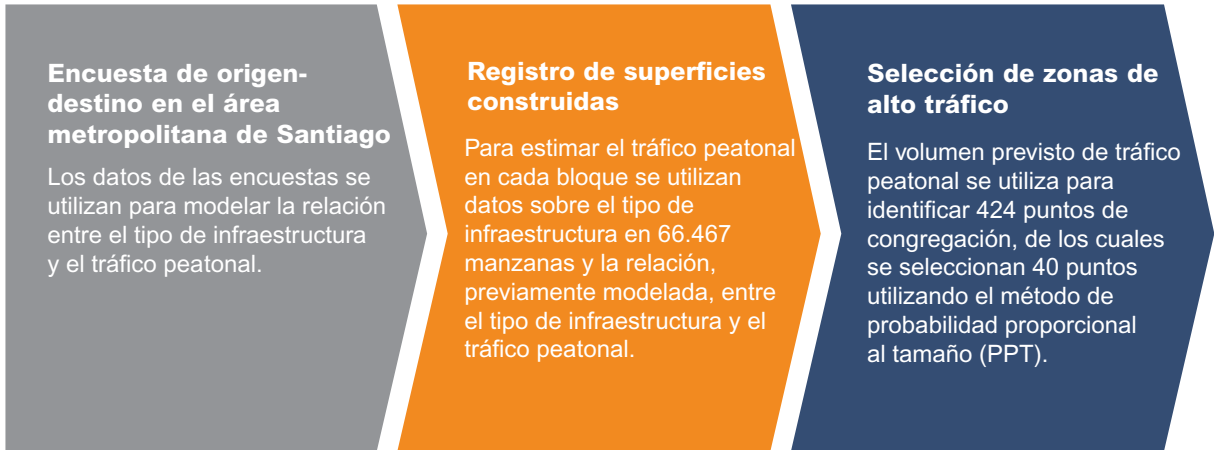
La noción básica del muestreo estratificado consiste en dividir toda la población en subpoblaciones (estratos) separadas y distintas. Se extraen muestras independientes dentro de cada estrato (Frankel, 1983). El muestreo estratificado garantiza que los grupos clave estén adecuadamente representados en la muestra. Además, los investigadores pueden querer estudiar ciertos subgrupos de población. El tamaño de la muestra para ese estrato puede aumentarse para asegurar un tamaño de muestra suficiente para el análisis de subgrupos separados. En la muestra completa, los estratos más grandes se pueden ponderar adecuadamente dentro de la muestra total (Frankel, 1983). Los estratos más grandes no necesitan ser alterados cuando se usan ponderaciones.

El muestreo estratificado es particularmente útil para la estimación de los cigarrillos ilícitos mediante el método de encuesta de análisis de cajetillas de fumadores. Las encuestas sobre el consumo de tabaco, como la Encuesta Mundial de Tabaquismo en Adultos (GATS, por sus siglas en inglés) (CDC, 2019), u otras encuestas nacionales, proporcionan información sobre la edad, el sexo y las características socioeconómicas de los fumadores. La información recopilada por la GATS puede utilizarse para informar a los estratos en otros proyectos de investigación. Por ejemplo, si las mujeres de 60 años o más representan el cinco por ciento de todos los fumadores del país, entonces los investigadores pueden usar el cinco por ciento para este subgrupo de edad y género en su estratificación.

El Proyecto de Evaluación de la Política Internacional de Control del Tabaco (Proyecto ITC) tiene algunos de los mejores ejemplos de muestreo estratificado en encuestas de fumadores (ITC, 2017). Las encuestas del Proyecto ITC, varias de las cuales son representativas a nivel nacional, estratifican a la población en consumidores y no consumidores de tabaco de forma conveniente. Los tamaños de las muestras, tanto de usuarios como de no usuarios, son representativos, de modo que cuando se combinan, el análisis de la encuesta arroja estimaciones precisas para toda la población del país. Sin embargo, el estrato de consumidores de tabaco está considerablemente aumentado (sobrerrepresentado). Esto permite a los investigadores realizar análisis adicionales con mayor precisión en este subgrupo. Debido al tamaño del estrato de fumadores, algunas encuestas del ITC son capaces de descomponer el estrato aún más. El subgrupo está clasificado por los consumidores de cigarrillos, los usuarios de otros productos de tabaco, así como los usuarios duales/mixtos (es decir, más de un producto de tabaco). El análisis se puede realizar a un nivel aún más granular. Las ponderaciones proporcionadas por la encuesta se pueden utilizar para combinar estratos de diferente tamaño a fin de que el análisis se realice en la muestra total (ITC, 2017).

El muestreo estratificado tiene dos advertencias importantes. En primer lugar, aunque el método garantiza que los subgrupos estén debidamente representados en la muestra, cada subgrupo (estrato) aún debe ser objeto de un muestreo adecuado. Por ejemplo, si se va a tomar una muestra de diez mujeres de 60 años o más, pero todas ellas se toman del mismo lugar, es probable que la muestra esté sesgada. Es probable que las personas que se encuentran en el mismo lugar compartan características similares (por ejemplo, ingresos, educación, estado civil, etc.). Para preservar la aleatoriedad dentro de la muestra, esta también debe ser controlada geográficamente. En segundo lugar, los investigadores deben ser conscientes de los posibles problemas que se derivan de una baja tasa de respuesta y de la estratificación. No sería estadísticamente apropiado, y por lo tanto sesgaría la muestra, acercarse a 100 mujeres de 60 años o más para recopilar información sobre solo 10 de ellas por desgaste. Por lo

Gráfico 3 Muestreo de los puntos de congregación usando el tráfico peatonal previsto en Santiago de Chile



general, con una baja cooperación, también hay importantes sesgos en las características de los sujetos (por ejemplo, los fumadores con mayores ingresos tienen menos probabilidades de mostrar sus cajetillas) (Sudman, 1983). Por consiguiente, en los lugares en que se espera una baja tasa de respuesta a la encuesta sobre el análisis de cajetillas de cigarrillos, el método de recolección de cajetillas desechadas puede ser una forma más apropiada para recopilar datos.

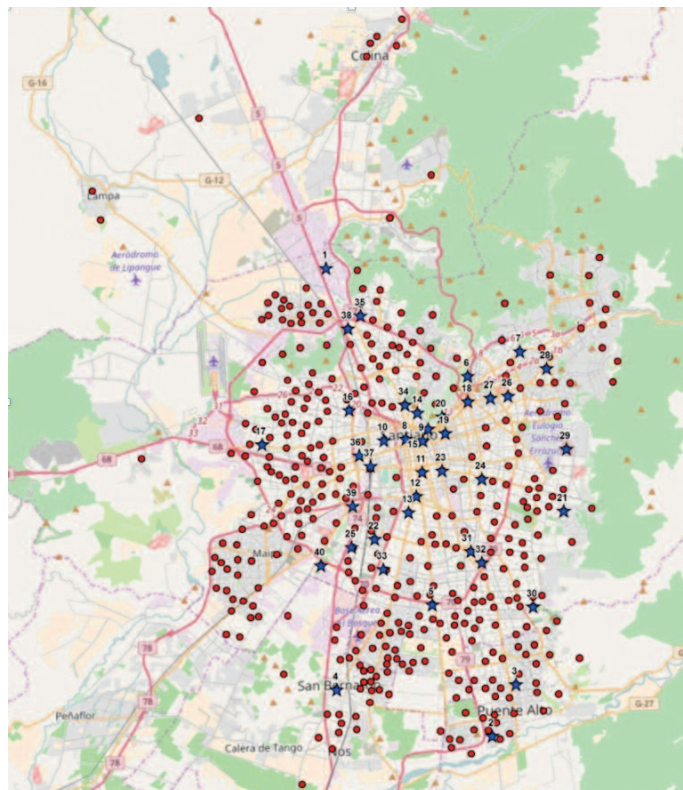
En un estudio sobre el comercio ilícito realizado en Santiago de Chile se utilizó una muestra estratificada y controlada geográficamente (Paraje *et ál.*, 2018). En el estudio se utilizaron las estimaciones de una encuesta nacional anterior sobre el consumo de drogas para estratificar su muestra por edad y sexo. Los investigadores recopilaron datos en zonas de alto tráfico. El estudio utilizó un proceso sofisticado para seleccionar estas áreas (Gráfico 3). La Sección 6.1 del Anexo proporciona más información sobre la selección del punto de congregación. El Gráfico 4 presenta la distribución geográfica de los 424 puntos de congregación (puntos rojos y estrellas azules), así como los 40 puntos de congregación seleccionados para la muestra (estrellas azules).

Un estudio sobre el comercio ilícito de cigarrillos en Colombia se aproximó de manera similar a una muestra estratificada por edad y género utilizando un muestreo geográficamente controlado (Maldonado *et ál.*, 2018). El estudio recopiló datos en las áreas que rodean a las instituciones educativas, estaciones de transporte público, parques de la ciudad, etc. Los puntos de estudio de la encuesta se distribuyeron en cada ciudad, de modo que las personas de los distintos distritos de la ciudad estuvieran igualmente representadas.

3.3.5.3 Muestreo por conglomerados

Como se ha señalado en secciones anteriores, el muestreo aleatorio simple se enfoca en elementos individuales seleccionados de toda la población. El muestreo estratificado divide a la población en subgrupos (estratos) basado en las características observables (por ejemplo, el sexo, la edad y la raza). El muestreo por conglomerados se encuentra entre una muestra aleatoria simple y una muestra estratificada. La población se divide en grupos (conglomerados). Se selecciona una muestra de esos grupos para el estudio. Se extrae una muestra aleatoria simple de los sujetos del estudio de cada conglomerado seleccionado. Un ejemplo de muestreo por conglomerados es un estudio que divide una ciudad en bloques (conglomerados) y luego toma muestras al azar de esos bloques. El análisis se realiza solo en los bloques muestreados.

Gráfico 4 Puntos de congregación utilizados en el estudio del comercio ilícito en el área metropolitana de Santiago de Chile



Origen: Mapa de los autores

Debido a la naturaleza del método de recolección de cajetillas desechadas, el muestreo por conglomerados es el enfoque de muestreo más apropiado. El muestreo aleatorio simple tradicional (es decir, la selección aleatoria de cajetillas de una lista o base de datos) sería casi imposible de aplicar. En el muestreo por conglomerados, la región de interés se divide en conglomerados de los cuales se escogen conglomerados aleatorios. Las cajetillas desechadas solo se pueden recolectar de los grupos elegidos. Hay varios ejemplos de estudios que utilizan la técnica de muestreo por conglomerados para el método de recolección de cajetillas desechadas: Merriman (2010), Stoklosa y Ross (2014), Barker *et al.* (2016), y Ross *et al.* (2019). Este método de muestreo se utilizó también en varias encuestas de análisis de cajetillas de fumadores (por ejemplo, Joossens *et al.*, 2014; Little *et al.*, 2019).

La primera etapa del muestreo de conglomerados con el método de recolección de cajetillas desechadas requiere que los investigadores seleccionen conglomerados geográficos. Los conglomerados se denominan unidades primarias de muestreo (UPM) y deben ser lo suficientemente pequeños. Durante la segunda etapa, se recoge una muestra de cajetillas en la UPM seleccionada. Stoklosa y Ross (2014) estimaron la prevalencia de la elusión y evasión fiscal a nivel ciudad en Varsovia (Polonia). Los investigadores seleccionaron una muestra de 30 UPM de 783 distritos electorales dentro de los límites de la ciudad. Barker y sus colegas (2016), que estimaron la prevalencia a nivel nacional de la elusión y evasión fiscal en los Estados Unidos, seleccionaron una muestra de 160 distritos escolares (UPM) de entre decenas de miles de distritos escolares del país. Además de los distritos electorales y los distritos escolares, otras divisiones geográficas, como las zonas censales, los códigos postales, las zonas y los distritos de las ciudades son posibles candidatos para las UPM en un estudio de recolección de basura.

Dos factores contribuyen a la decisión sobre qué división geográfica se utilizará para una UPM: 1) La información disponible para cada división geográfica; y 2) el tamaño de la división geográfica. El proceso de muestreo requiere información de los fumadores, como dónde viven, fuman y desechan sus cajetillas. Esta es una información vital para asegurar una muestra representativa. Los mejores candidatos para la UPM son aquellos para los que esta información está disponible. Es importante señalar que una zona censal más pequeña tendrá menor varianza. Se prefieren las UPM más pequeñas, excepto en los casos en los que visitar varias UPM pequeñas sea

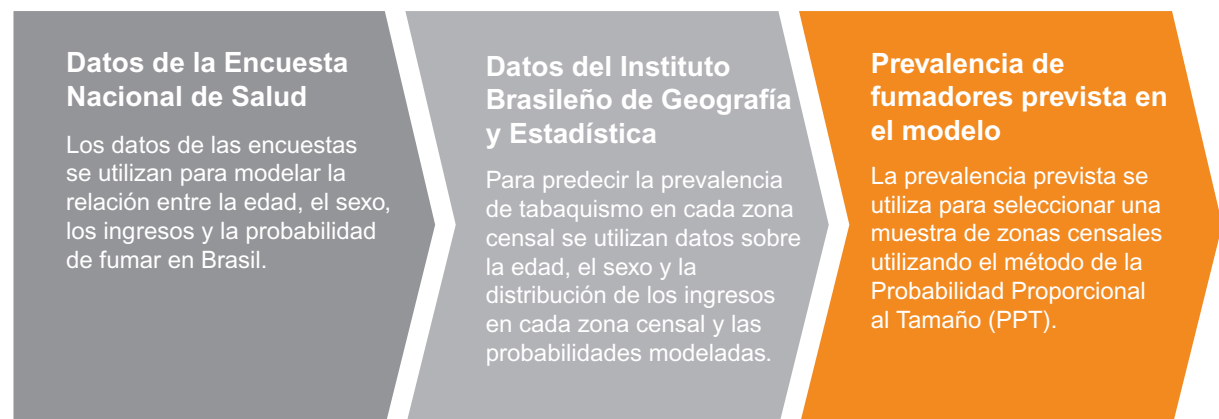
logísticamente difícil o costoso. Como alternativa, los investigadores deberían considerar varias UPM de mayor tamaño y recopilar más datos de cada una de ellas.

Aunque teóricamente se desearía información sobre el consumo de productos de tabaco y el comportamiento de desecho para cada UPM, en la práctica no suele estar disponible. En este caso, los comportamientos de fumar y desechar deben estimarse usando la información existente de la UPM. Los investigadores pueden utilizar información preexistente sobre las características socioeconómicas de cada población de la UPM. Estas estimaciones pueden utilizarse para inferir el número de fumadores en cada unidad. Las características de sexo, edad e ingresos son los mejores predictores del consumo de tabaco. En general, los hombres son más propensos a fumar que las mujeres. En cuanto a la edad, la probabilidad de fumar tiene forma de campana: al principio, las tasas de fumadores aumentan con la edad hasta que alcanzan su punto máximo, después de lo cual disminuyen. La prevalencia de tabaquismo también tienen una fuerte correlación con los ingresos. El que la prevalencia de tabaquismo aumenten o disminuyan con los ingresos depende de la epidemia de tabaquismo del país o de su etapa de desarrollo económico y educativo.

Un estudio brasileño estimó la población fumadora en cada UPM para un método de recolección de cajetillas desechadas. El estudio tenía por objeto estimar la elusión y evasión fiscal del tabaco en las ciudades de Río de Janeiro, Sao Paulo, Belo Horizonte, Campo Grande y Joao Pessoa. Los investigadores utilizaron zonas censales como UPM. Para seleccionar las mejores muestras basadas en el censo, los investigadores necesitaban evaluar el número de fumadores en cada UPM. El Instituto Brasileño de Geografía y Estadística (IBGE), que realiza el censo, proporcionó información sobre la edad, el sexo y la distribución de los ingresos de los residentes en cada zona censal. Estos datos se combinaron con los datos de la Encuesta Nacional de Salud sobre las características de los fumadores para estimar el número de fumadores en cada zona censal. La edad, el género y los ingresos se utilizaron para predecir la probabilidad de fumar cigarrillos en cada zona censal, utilizando un modelo de regresión logística para cada zona. Las probabilidades de fumar previstas se utilizaron para determinar el número de fumadores (Gráfico 5) (Carvalho Figueiredo *et ál.*, próxima publicación). El Anexo 6.4 proporciona más detalles sobre este estudio.

El número de UPM incluidas depende del tamaño de la muestra final deseada. El estudio descrito en el Capítulo 3.3.4 estima que el número deseado por ciudad mexicana es de 670 cajetillas, lo que requiere que los recolectores caminen aproximadamente 52 millas (83 kilómetros). Las Áreas Geoestadísticas Básicas, o AGEB, fueron seleccionadas como UPM. Los investigadores caminaron a lo largo de todas las rutas principales en cada AGEB. Dado que la longitud media de las rutas en cada AGEB era de aproximadamente 5 millas (8 kilómetros),

Gráfico 5 Muestreo de regiones geográficas mediante el método de la prevalencia prevista del tabaquismo y la Probabilidad Proporcional al Tamaño (PPT) en Brasil



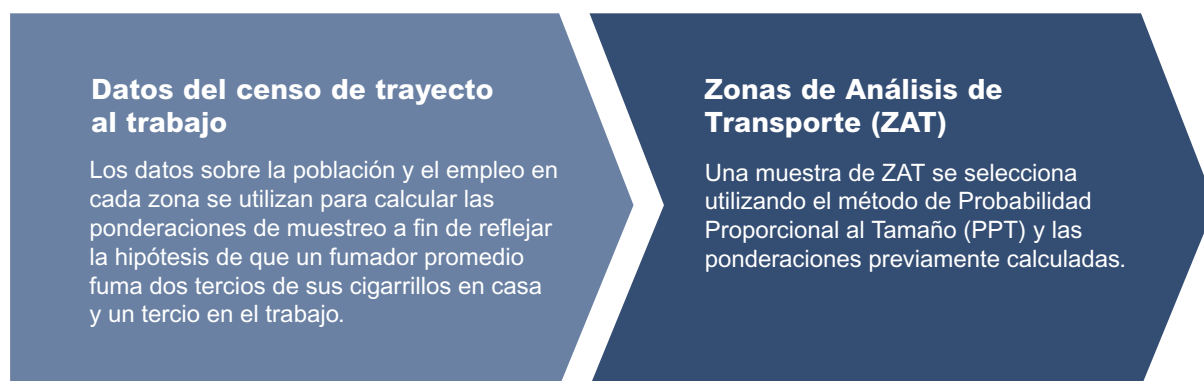
los investigadores determinaron que necesitaban visitar en promedio 11 AGEB por ciudad. El número de AGEB oscilaba entre 9 y 14 y la variabilidad dependía de la densidad de calles en cada ciudad (Sáenz de Miera Juárez *et ál.*, 2020).

Una vez que se determina el número de elementos de estudio en cada UPM (conglomerado) y se elige el número de UPM, la selección de las UPM puede ocurrir de dos maneras. Si se espera que el tamaño de la población de cajetillas en cada UPM sea similar entre todas las unidades, la muestra de las UPM se puede seleccionar al azar. Sin embargo, si el tamaño de la población de las cajetillas difiere significativamente entre las UPM, las UPM se deben seleccionar utilizando la Probabilidad Proporcional al Tamaño (PPT). El muestreo simple aleatorio de las UPM es sencillo. El Anexo 6.5 provee un ejemplo de selección aleatoria de UPM usando un código de Stata. El muestreo de PPT es solo un poco más complicado. A cada UPM se le da una ponderación apropiada basada en el número de fumadores y/o en la estimación de cajetilla desechadas. Las UPM se muestrean entonces en base a las ponderaciones. El Anexo 6.4 tiene el código para la selección de PPT del mencionado estudio brasileño.

Un muestreo PPT con dos o más variables de ponderación es más complejo. En un estudio realizado en Chicago (Gráfico 6) se utilizó ese enfoque para estimar la elusión y la evasión fiscal de los cigarrillos (Merriman, 2010). Las zonas de transporte se utilizaron como UPM. No se disponía de información sobre el comportamiento de los fumadores en las zonas de transporte, por lo que los investigadores tenían que estimar el número de fumadores utilizando los datos de las características disponibles. Utilizaron la población y el número de personas que trabajan en cada área como predictores. A las zonas con mayor población se les dio una ponderación del 100%. Se dio una ponderación del 49% a las regiones con mayor empleo (Gráfico 6).² Las ponderaciones utilizadas reflejan la suposición de que un fumador medio fuma dos tercios de sus cigarrillos en casa y un tercio en el trabajo. El objetivo era dar cuenta de los fumadores que tiran sus cajetillas cerca de su residencia así como de su lugar de trabajo. Las zonas de transporte se ponderaron utilizando el método PPT para seleccionar una muestra de las zonas donde se recolectaron las cajetillas. En el Anexo 6.6 se presenta un código de Stata que se puede utilizar para este tipo de muestreo.

Una distinción importante entre el estudio brasileño y el de Chicago es que se asumió que todos en Chicago tenían la misma probabilidad de ser fumadores. El estudio de Chicago no hizo ninguna distinción entre la edad, el sexo, los ingresos u otras características observables en sus probabilidades de fumar.

Gráfico 6 Muestreo usando el método de Probabilidad Proporcional al Tamaño (PPT) en Chicago



² Merriman usa un ponderador de 49% y no de 50% por una restricción técnica. El empleo está altamente concentrado en algunas zonas de Chicago. El ponderador más alto permitido en la selección de 100 ZATs de un total del 930 ZATs de Chicago sin reemplazo fue de 49.

3.3.6 Etapas de muestreo

Solo en el caso de un muestreo aleatorio simple el diseño de estudio es un muestreo de una sola etapa. Todos los demás métodos de muestreo requieren un muestreo de varias etapas. Un enfoque de muestreo de varias etapas se refiere a la extracción de muestras subsecuentes en unidades de muestreo cada vez más pequeñas. Los estudios mencionados en las secciones anteriores ofrecen buenos ejemplos. En un estudio chileno sobre el comercio ilícito realizado en Santiago de Chile se seleccionaron 40 puntos de congregación (primera etapa), tras lo cual se seleccionó a los fumadores mediante el método de muestreo estratificado (segunda etapa) (Paraje *et ál.*, 2018). De manera similar, el estudio mexicano seleccionó en promedio 11 AGEB por ciudad (primera etapa), después de lo cual todas las cajetillas de cigarrillos desechadas se recolectaron a lo largo de las calles principales (segunda etapa) (Sáenz de Miera Juárez *et ál.*, 2020).

A veces es necesario utilizar más de dos etapas de muestreo. Por ejemplo, en un estudio en Georgia sobre el comercio ilícito se seleccionaron las UPM utilizando el método PPT (primera etapa). Posteriormente, los hogares se seleccionaron dentro de cada UPM (segunda etapa). Por último, se seleccionó a los fumadores al azar dentro de cada hogar utilizando el método de la cuadrícula de Kish (tercera etapa) (Little *et ál.*, 2019).

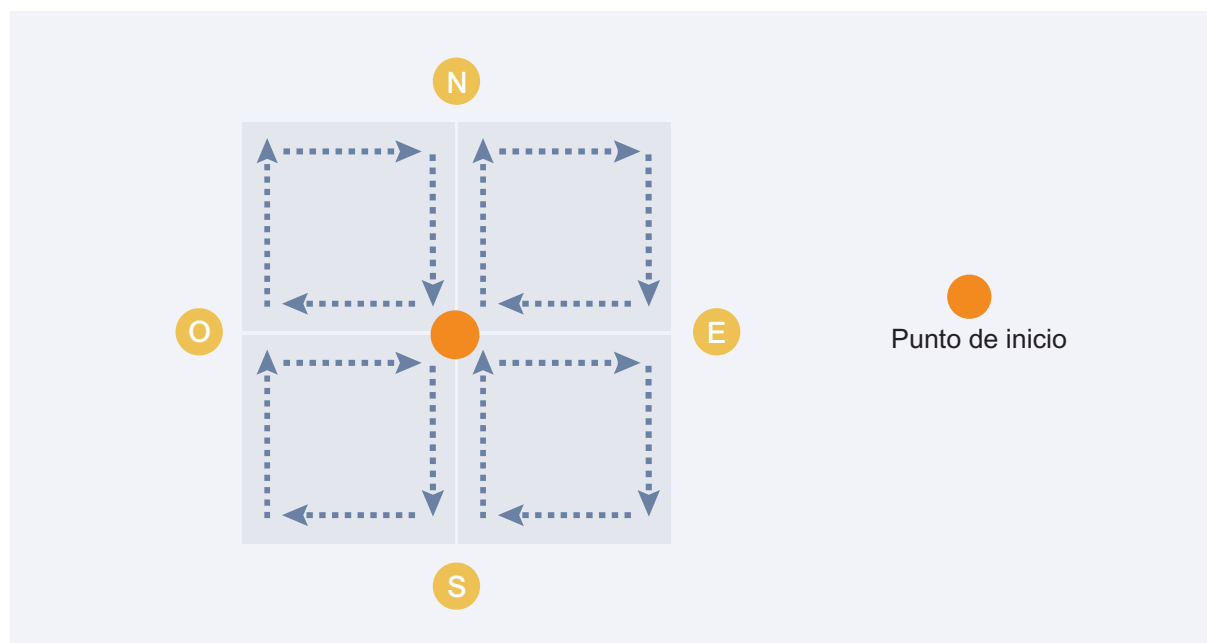
Para las recolecciones de cajetillas desechadas, casi siempre es necesario utilizar más de dos etapas de muestreo. Rara vez es factible recolectar todas las cajetillas desechadas dentro de una UPM. Cuando las UPM son grandes, los investigadores necesitan designar rutas (segunda etapa) donde se recolectarán las cajetillas desechadas. Las cajetillas se recolectarán sistemáticamente solo de partes de la UPM. La longitud de la ruta está determinada por varios factores: el número de cajetillas que se espera encontrar por unidad de distancia (millas o kilómetros), el número de UPM seleccionadas en la muestra y el tamaño total deseado de la muestra de cajetillas.

Un protocolo de caminata para la recolección de cajetillas asegurará que se tomen las muestras adecuadas. El protocolo debe cumplirse estrictamente durante el proceso de recolección de datos en todas las UPM. Un estudio sobre el comercio ilícito realizado en Buenos Aires, Argentina, llevado a cabo por Pizarro *et ál.*, implementó el siguiente protocolo de caminata: los encuestadores caminan 100 metros hacia el norte desde el punto de partida. Cada 100 metros, el recolector gira a la derecha hasta completar un cuadrado con perímetro de 400 metros. Este patrón se repite cuatro veces, cada vez eligiendo una dirección de inicio diferente. El resultado final es una cuadrícula de dos por dos de 400 metros cuadrados (Gráfico 7). Se recolectan todas las cajetillas desechadas a lo largo de la ruta (Pizarro *et ál.*, en revisión). La muestra es imparcial cuando los recolectores recorren las rutas preestablecidas. La mejor práctica es seleccionar todas las cajetillas a lo largo de las rutas preestablecidas. Los diferentes números de cajetillas recolectadas por UPM reflejan con mayor precisión el número de fumadores, así como la intensidad del consumo de tabaco, que no se puede medir con la recolección por cuotas.

Los investigadores también deben considerar el punto de partida apropiado en el escenario descrito anteriormente. Como el punto de partida está en el epicentro del cuadrado mayor de 1.600 metros, la aleatoriedad de la ubicación del punto de partida es crucial. Escoger a mano un punto de partida no garantiza una representación exacta de la población total de fumadores dentro de una UPM. Por ejemplo, si las cajetillas ilícitas son más frecuentes alrededor de las plazas del mercado, las recolecciones de cajetillas que siempre comienzan en la plaza del mercado representarán en exceso las cajetillas de cigarrillos ilícitas. La elección de un punto de partida aleatorio para cada ruta de recopilación de datos garantiza que las cajetillas recolectadas representen mejor las cajetillas que se suelen fumar en una zona determinada.

Las ubicaciones de inicio se pueden elegir de varias maneras. En los estudios en que las UPM se basan en zonas censales u otras divisiones geográficas oficiales, suele existir una lista de direcciones físicas dentro de cada una de esas UPM. Una de estas direcciones se puede seleccionar al azar de la lista para cada UPM. Alternativamente, un punto de partida se puede seleccionar espacialmente al azar. Si existe un mapa digital (por ejemplo, *shapefiles*), los puntos de partida de la UPM se pueden seleccionar al azar. En el programa ArcGIS (un software de sistemas de información geográfica), se puede utilizar la función de "Crear puntos aleatorios". Los recolectores pueden usar las coordenadas GPS para localizar sus puntos de partida. El Gráfico 8 a continuación presenta los puntos de partida seleccionados al azar en las zonas censales en uno de los barrios de Río de Janeiro, Brasil (Carvalho Figueiredo, próxima publicación).

Gráfico 7 Una ruta de 1.600 metros recorrida en cada una de las zonas censales seleccionadas en el estudio de cajetillas ilícitas en Argentina



Origen: Pizarro *et ál.*, en revisión

Aparte de ArcGIS, hay varios generadores de puntos aleatorios gratuitos en línea. Estos pueden generar puntos aleatorios dentro de una región circular (por ejemplo, centrada en el medio de la UPM) o dentro de una región rectangular (por ejemplo, definida como los límites norte, sur, oeste y este de cada UPM). Aunque cualquier proyección de mapa distorsiona la forma de grandes áreas (por ejemplo, continentes y países), las UPM investigadas en las recolecciones de cajetillas desechadas son siempre mucho más pequeñas y no se ven afectadas significativamente por la proyección de mapas. Por lo tanto, la selección de puntos de partida aleatorios en el mapa debería funcionar en una recolección de cajetillas desechadas.

Al elegir el punto de partida del estudio, es imperativo no dirigirse a los lugares donde es más probable que se encuentren cajetillas de productos de tabaco ilícitos. De hecho, las recolecciones de cajetillas desechadas pagadas por la industria tabacalera suelen ser criticadas por centrarse en áreas en las que es más probable el consumo de productos que no cumplen con los impuestos (Rowell *et ál.*, 2014). Las encuestas de la industria en el Reino Unido se llevan a cabo en eventos deportivos, donde las cajetillas extranjeras/ilícitas tienen una mayor probabilidad de consumirse (Gilmore *et ál.*, 2014). De manera similar, un estudio de recolección de cajetillas financiado por la industria en Alemania sistemáticamente sobrerrepresentó las regiones geográficas a lo largo de la frontera del país y alrededor de las bases militares de los Estados Unidos, donde se espera que haya más cigarrillos extranjeros (Adams *et ál.*, 2011). Para medir objetivamente la elusión y la evasión fiscal del tabaco es necesario reunir datos de una muestra aleatoria de todos los cigarrillos consumidos.

Los investigadores no deben enviar a los encuestadores a zonas con altos porcentajes de consumidores ilícitos de cigarrillos, a menos que ello forme parte de un diseño de estudio para reunir información de esos puntos de conflicto. En cambio, el enfoque debe ser en todos los cigarrillos que se consumen. Si un local con una alta incidencia de cigarrillos ilícitos cae en la ruta de los recolectores, entonces debe formar parte de la recopilación de datos. Del mismo modo, el estudio no debe centrarse únicamente en las zonas de altos ingresos. El objetivo es no tomar muestras excesivas o insuficientes de ciertas zonas debido a nociones preconcebidas, ya que esto introducirá un sesgo en el análisis y hará que los resultados no sean confiables.

Gráfico 8 Puntos seleccionados al azar en cada zona censal en el barrio de Ramos de Río de Janeiro, Brasil



Origen: El mapa de los autores fue creado usando archivos de formas disponibles públicamente y el programa ArcGIS

3.4 Recopilación de datos

3.4.1 Construcción del cuestionario

El desarrollo de instrumentos para la encuesta de análisis de cajetillas de fumadores es relativamente sencillo. Para empezar, el encuestador debe recitar una breve explicación y obtener el consentimiento del encuestado para ser entrevistado. La explicación debe incluir información sobre el objetivo del estudio y los derechos del participante. Una vez dado el consentimiento, las primeras preguntas de una encuesta deben ser preguntas de selección. Estas están diseñadas para determinar si el sujeto es elegible para participar en el estudio. Las preguntas de selección deben incluir los criterios de inclusión en el estudio, como la edad del participante o su condición de fumador. Las encuestas de muestreo estratificado deben incluir también información de los participantes sobre el género, los ingresos, etc.

A continuación, el encuestador debe pedirles a los fumadores que muestren su cajetilla de cigarrillos. Si los fumadores disponen de varias cajetillas de cigarrillos (por ejemplo, que haya varias cajetillas en el hogar), los lineamientos de investigación deben determinar cómo debe proceder el encuestador. Una práctica óptima sería incluir solo la última cajetilla comprada (por ejemplo, Maldonado 2018) o, si se compran varias cajetillas, la cajetilla en la que se fumó el último cigarrillo (por ejemplo, Paraje 2018). De esta forma, no se dirigen específicamente a determinadas marcas o cigarrillos ilícitos. Incluir algunas cajetillas de cigarrillos y no otras puede llevar a una muestra sesgada. El hecho de representar solo la cajetilla de cigarrillos que se haya comprado o fumado más

recientemente no discrimina subjetivamente y refleja con mayor precisión la conducta de fumar. El objetivo es tener una muestra imparcial para proporcionar la representación más precisa del comportamiento de la población.

El siguiente ejemplo hipotético sirve para demostrar el beneficio de pedir la última cajetilla comprada. Supongamos que una muestra consta de cien fumadores, que fuman las marcas X y Z. Todos los fumadores compran cualquiera de las dos marcas y alternan su comportamiento de compra entre X y Z. Cada fumador compra una cajetilla de X y posteriormente compra dos cajetillas de Z. La marca X es la preferida, pero la marca Z es más barata porque es la marca ilícita y por lo tanto se compra con mayor frecuencia. Cada fumador tiene al menos una cajetilla de cualquiera de las dos marcas de cigarrillos. Si un encuestador le pide al consumidor que muestre solo su marca preferida, los cien miembros mostrarían los cigarrillos de la marca X. En consecuencia, la prevalencia estimada de cigarrillos ilícitos sería cero. Por otra parte, si el encuestador le pregunta a cada fumador si tiene cigarrillos ilícitos, entonces la prevalencia estimada sería del 100 % para la población. Ninguno de los dos escenarios capta el verdadero comportamiento de los fumadores en la población. Sin embargo, una encuesta en la que se pida a cada fumador que muestre su última cajetilla comprada, llegará a una prevalencia de cigarrillos ilícitos del 67%, que es la verdadera tasa de compra de la marca ilícita Z.

El estudio debe captar con precisión el índice de tabaquismo ilícito. Esto se conoce como la intensidad con la que fuma el consumidor. Una pregunta importante para captar la intensidad con la que fuma el consumidor es cuántos cigarrillos se fuma por día o por semana. Sin esta información es imposible estimar la prevalencia de los cigarrillos ilícitos. El siguiente ejemplo hipotético sirve para demostrar la necesidad de preguntar sobre la intensidad con la que fuma el consumidor. Supongamos que hay diez fumadores, cada uno de los cuales fuma solo una marca. Ocho de los diez fumadores fuman una marca legal X a razón de cinco cigarrillos por día. Los otros dos fumadores fuman una marca ilegal Z a razón de 20 cigarrillos por día (una cajetilla completa). Dado que las tasas de fumadores son diferentes entre los dos tipos de consumidores, si en una encuesta no se incluyera una pregunta sobre el índice de tabaquismo, no se podría describir con exactitud la prevalencia de cigarrillos ilícitos en el grupo. La prevalencia correcta de cigarrillos ilícitos entre los diez fumadores es del 50% (por cada 40 cigarrillos de la marca X que fuman los ocho fumadores, hay otros 40 cigarrillos de la marca Z que fuman los dos fumadores). Así pues, la verdadera prevalencia de los cigarrillos ilícitos solo se puede captar mediante la inclusión de ponderaciones apropiadas sobre la intensidad con la que fuman los participantes.

Para disminuir la tasa de desgaste, los encuestadores deben explicar el propósito de la encuesta. A fin de que los encuestados comprendan por qué deben dedicar tiempo a la encuesta, los encuestadores deben subrayar que la encuesta se realiza para mejorar la salud pública. Si la encuesta se realiza en colaboración con una organización de salud pública, debe mencionarse el nombre de la organización. Las personas de todos los grupos demográficos generalmente conocen y confían en las organizaciones de salud pública y, por lo tanto, la mención del nombre de la organización aumentará la probabilidad de respuesta.

Una vez que se muestra la última cajetilla comprada, el encuestador necesita recopilar datos sobre las características de ésta. La encuesta debe incluir todas las preguntas necesarias sobre las características de la cajetilla que ayuden a los investigadores a distinguir entre las cajetillas que cumplen y las que no cumplen con las regulaciones (véase la Sección 3.2.2). Las preguntas deben centrarse en la marca de la cajetilla y la presencia de un sello fiscal apropiado, así como en la presencia, el tipo, el tamaño y el lenguaje de la advertencia sanitaria, y otros marcados de la cajetilla requeridas por la ley.

La información será más objetiva si se recopila mediante la observación de la cajetilla durante la encuesta o mediante la obtención de una cajetilla del encuestado (por ejemplo, para un futuro examen pericial) en lugar de que el encuestado responda a las preguntas. Una buena práctica es tomar fotos de cada lado de la cajetilla presentada. Estas imágenes pueden ayudar a identificar los errores de codificación, ya que los investigadores pueden remitirse a las imágenes de la cajetilla si surgen problemas. Estas imágenes también pueden servir como documentación de estudio en caso de que se cuestionen los resultados de éste. Las fotografías tomadas con un teléfono móvil tendrán la fecha, la hora y el lugar apropiados (coordenadas GPS) asociados a cada fotografía. Esto facilita el cotejo con el estudio apropiado, la facilidad de búsqueda y es una información útil durante la fase de análisis de los datos.

La opción múltiple es el mejor formato para las preguntas de la encuesta porque estandariza las respuestas, mitigando posibles errores de codificación, y ayuda al encuestador a completar la encuesta rápidamente. Por ejemplo, la pregunta sobre la marca de cigarrillos debe incluir una lista de las marcas de cigarrillos disponibles en el mercado. La opción de introducir la marca manualmente también debe estar presente en caso de que la marca no esté en la lista. Por último, si la encuesta tiene una pregunta de sellos fiscales, entonces la pregunta se debe formular cuidadosamente. En algunos países, los sellos fiscales en el empaque exterior o en el celofán se arrancan al abrir la cajetilla. Esta característica está diseñada para evitar la reutilización de los sellos fiscales. La pregunta sobre los sellos fiscales debe formularse de manera que pueda distinguir entre las cajetillas con residuos de sellos fiscales y las cajetillas sin sello fiscal (por ejemplo, los investigadores de Mongolia buscaron residuos de pegamento que indicaran que el sello se había quitado [Ross *et ál.*, 2019]).

Dos datos pueden ser suficientes para estimar la prevalencia de las cajetillas de cigarrillos que no cumplen con las normas. Esta información es la intensidad de consumo de tabaco y las características de la última cajetilla de cigarrillos comprada. Sin embargo, hay tres excepciones en las que se puede requerir información adicional. En primer lugar, cuando una parte significativa de los encuestados no puede o no quiere mostrar sus cajetillas de cigarrillos, los investigadores podrían querer hacer un seguimiento con varias preguntas adicionales. Como se ha examinado en secciones anteriores, la incapacidad (por ejemplo, los fumadores de cigarrillos individuales) o la falta de voluntad (por ejemplo, las normas culturales) para mostrar las cajetillas de cigarrillos probablemente esté correlacionada con las características demográficas y socioeconómicas del fumador. La eliminación de estas personas de la muestra introduce un sesgo en la estimación del comercio ilícito. Las preguntas adicionales formuladas se pueden utilizar más adelante para imputar los datos que faltan sobre el cumplimiento de las obligaciones fiscales. Las preguntas adicionales deben incluir información sobre la marca, la cantidad y el precio del último cigarrillo comprado, así como la ubicación de la última compra y si el encuestado fuma cajetillas o cigarrillos sueltos. Al entrevistar a un fumador de cigarrillos individuales, el encuestador debe pedirle que muestre el cigarrillo, ya que suele tener la marca impresa.

La segunda circunstancia que requiere información adicional sobre los fumadores es cuando se necesita un análisis a fondo del comercio ilícito. Por ejemplo, la recopilación de datos sobre la educación, los ingresos y la ocupación del fumador podría ayudar a determinar si ciertos grupos socioeconómicos u ocupaciones son más propensos a fumar cigarrillos ilícitos. Incluir una pregunta sobre los lugares de compra podría ayudar a comprender la cadena de suministro de los cigarrillos que no cumplen con las normas. Asimismo, la información sobre el tipo de establecimiento de venta al público en el que se compraron los cigarrillos es útil. Distinguir los cigarrillos comprados durante un viaje al extranjero, en una tienda libre de impuestos o en un mercado local ayuda a distinguir la elusión y la evasión fiscal.

En tercer lugar, puede ser necesario obtener cajetillas de los fumadores si se requiere un examen pericial adicional de la cajetilla o del sello fiscal. En esos casos, se les podría dar a los fumadores una pequeña remuneración por proporcionar su cajetilla a los encuestadores.

Por supuesto, existe un compromiso entre la recopilación de más información y el tiempo necesario para completar la encuesta. Si bien la recopilación de más información ayudará a obtener mejores estimaciones, el aumento del número de preguntas podría desincentivar a las personas a participar en la encuesta. Los investigadores deben sopesar la información adicional obtenida frente a la tasa de desgaste potencialmente más alta.

En el Anexo 6.2 se presenta un cuestionario de muestra a modo de orientación. El cuestionario se debe adaptar a las necesidades de cada estudio, teniendo en cuenta las cuestiones examinadas en el presente capítulo. Lo más importante es que la encuesta debe cumplir con las normas éticas locales. También se invita a los investigadores a que se pongan en contacto con los autores de las investigaciones que se examinan en este conjunto de herramientas que utilizaron encuestas (por ejemplo, Argentina [Prizarro *et ál.*, en revisión], Brasil [Carvalho Figueiredo, próxima publicación], Chile [Paraje *et ál.*, 2018], Colombia [Maldonado *et ál.*, 2018], Georgia [Little *et ál.*, 2019], México [Sáenz de Miera Juárez, 2020], Mongolia [Ross *et ál.*, 2019], Polonia [Stoklosa y Ross, 2014] y seis países del Sureste de Europa: Albania, Bosnia y Herzegovina, Kosovo, Montenegro, Macedonia del Norte y Serbia [Zubovic *et ál.*, próxima publicación]). Los autores de este conjunto de herramientas pueden facilitar el contacto si es necesario.

3.4.2 Aprobación ética

La encuesta debe aprobarse por el Consejo de Revisión Institucional (IRB, por sus siglas en inglés). Es muy probable que el método de recolección de cajetillas desechadas no requiera la aprobación del IRB porque los sujetos humanos no están directamente involucrados. Sin embargo, algunos países pueden requerir un examen ético para evaluar los riesgos a los que se exponen los investigadores. Cuando se trata de sujetos humanos, se requiere la aprobación de el IRB. Aunque la encuesta sobre el análisis de cajetillas de fumadores se centra en la cajetilla de cigarrillos, se requiere la aprobación de el Consejo de Investigación Institucional porque se trata de sujetos humanos.

El IRB necesita dar su autorización antes de que comience la recopilación de datos. Lo ideal es que la autorización de la IRB se obtenga en el país donde se realiza el estudio. El proceso de autorización puede llevar varias semanas. Las encuestas de análisis de cajetillas deben incluir tiempo adicional para la autorización del IRB en el cronograma del proyecto. Una consideración importante es que, si bien la venta de productos de tabaco ilícitos puede ser ilegal, el consumo de esos productos de tabaco tiende a no serlo. Cuando los fumadores no participan en una actividad ilegal, es posible que los investigadores obtengan una exención de el IRB. En el Anexo 6.3 se muestra una solicitud de exención del IRB. Es necesario enviar un paquete completo al IRB para su revisión, incluyendo la documentación completa del estudio, la metodología de muestreo, el instrumento de la encuesta, el acuerdo con la empresa entrevistadora, etc. Si no se obtiene la autorización adecuada de el IRB, es posible que el estudio sea no publicable en revistas de prestigio revisadas por pares y no se podrán presentar los resultados en el ámbito académico.

3.4.3 Ejecución de la encuesta

El objetivo del proceso de recopilación de datos es reunir la información de manera coherente para que se pueda reproducir. Si bien la realización del proceso de recopilación de datos para estas encuestas en la propia empresa es manejable, delegar el proceso de recopilación de datos a una empresa de estudios de mercados es una alternativa viable. Los grupos de investigación que disponen de tiempo limitado, recursos humanos escasos y fondos suficientes pueden preferir la opción de la contratación externa.

Existen varias empresas internacionales de investigación de mercados muy conocidas, así como encuestadores nacionales más pequeños. Este conjunto de herramientas no respalda a ninguna empresa en particular. Como pauta general, es crucial que la empresa de investigación de mercados esté dispuesta a trabajar bien con el equipo de investigación. Las empresas de investigación de mercados se especializan en trabajar con pequeñas y medianas empresas. Estas empresas suelen carecer de la capacidad de realizar estudios de investigación y análisis por su cuenta y, por lo general, buscan un producto acabado que incluya un informe o la presentación de las conclusiones del estudio. Las empresas de investigación de mercado generalmente no están familiarizadas con los clientes que participan en el estudio, incluso en la etapa de diseño, la construcción de la encuesta y el análisis de los datos en bruto. La empresa de investigación de mercados debe saber de antemano que los investigadores participarán activamente en cada etapa del estudio.

La tarea asignada a la empresa de investigación de mercados debe estar claramente definida y esbozada, ya que los grupos de investigación son un cliente atípico para ellos. Sin embargo, se ha hecho antes con éxito. Recientemente, una empresa de estudios de mercados realizó una encuesta en seis países de la región del Sureste de Europa, denominada Encuesta sobre el Consumo de Tabaco en el Sureste de Europa (STC-SEE, por sus siglas en inglés) (Zubovic *et ál.*, próxima publicación). Un consorcio de centros de estudio de la región, en asociación con el equipo de Tobacconomics de la University of Illinois at Chicago, elaboró el cuestionario para la STC-SEE, que incluía una encuesta de análisis de cajetillas de fumadores, y luego analizó los datos recopilados por la empresa. Se definieron claramente las funciones de la empresa de estudios de mercados y de los asociados del consorcio de investigación, y los datos de la encuesta se están utilizando en el análisis de la prevalencia de consumo de tabaco y la elusión y evasión fiscal en la región.

Tanto si el proceso de recopilación de datos se realiza en la propia empresa como si se subcontrata, es esencial realizar un estudio piloto para poner a prueba los instrumentos de la encuesta. Como ya se ha mencionado, los pilotos pueden ser útiles para aproximar la tasa de respuesta y, en el caso de las recolecciones de cajetillas

desechadas, el número de cajetillas desecharas que se encuentran a lo largo de una ruta predefinida. El piloto también sirve para proporcionar una estimación de la duración de la entrevista e identificar posibles problemas con la encuesta. Como mínimo, los investigadores deberían pedir información a otros investigadores que no participan en el proyecto.

El momento de la recopilación de datos requiere una consideración adicional. En el caso de la recolección de cajetillas desecharas, el momento de la recopilación de datos es de suma importancia. Las calles se limpian regularmente, quitando las cajetillas desecharas de las calles. La frecuencia y el momento en que se limpian las calles será una información necesaria para que los investigadores la tomen en consideración. Para las encuestas de análisis de cajetillas de fumadores, es necesario tener en cuenta el momento en que se realiza la encuesta en lugares públicos o en las visitas a los hogares. Diferentes escenarios se verán afectados por diferentes tiempos.

Antes de poner en práctica el proceso de recopilación de datos, se requiere cierto trabajo administrativo. Para el método de recolección de cajetillas desecharas, los investigadores necesitan crear rutas de caminata en mapas para los recolectores. Es necesario organizar otro tipo de equipo, como bolsas de plástico con un número de identificación único y guantes. Los recolectores deben utilizar las mejores prácticas y tomar una fotografía de la cajetilla desechara y de la bolsa de recolección con el número de identificación. Alternativamente, el número de identificación de la bolsa se puede marcar en el mapa de caminata. Aunque puedan parecer tediosas, estas preparaciones ayudan a los investigadores a identificar el lugar de recolección de la cajetilla, lo que puede ser relevante más tarde. Además, es preciso prestar atención a la coordinación entre los lugares de recolección de cajetillas y los centros de investigación, especialmente en los casos en que las cajetillas se recolecten en grandes regiones. Por último, el almacenamiento de las cajetillas de cigarrillos debe estar asegurado. Aunque el almacenamiento de cajetillas vacías puede ser costoso y problemático, una buena práctica consiste en mantener las cajetillas recolectadas el mayor tiempo posible, en caso de que se cuestionen los resultados del estudio o en caso de que se necesite un análisis adicional de las cajetillas.

Para las encuestas de análisis de cajetillas de fumadores, los encuestadores deben recibir el protocolo de la encuesta, con la información sobre el procedimiento de recopilación de datos, un manual para el encuestador que refuerce las habilidades aprendidas durante la capacitación básica y los materiales de referencia necesarios para que los encuestadores completen el análisis de cajetillas (por ejemplo, en el paquete de información debe incluirse una copia impresa de todas las advertencias sanitarias aprobadas para ayudar al encuestador a distinguir entre las advertencias sanitarias nacionales y las extranjeras). Tanto para la recolección de cajetillas desecharas como para la encuesta de análisis de cajetillas de fumadores, los encuestadores deben tener acceso a un teléfono celular funcional con suficiente carga de batería y almacenamiento de fotografías disponible.

Es aconsejable que los investigadores estén al alcance durante el proceso de capacitación, incluso cuando la recopilación de datos se subcontrate a empresas de estudio de mercados. Los investigadores que están íntimamente involucrados en los detalles del estudio de investigación también se benefician del proceso de capacitación. Los investigadores pueden transmitir la importancia del estudio e inculcar el mismo entusiasmo a los encuestadores. Proporcionarles a los encuestadores el contexto del estudio puede aumentar la motivación y garantizar un personal bien capacitado. La presencia de los investigadores garantiza también la difusión de la información y los materiales necesarios.

La buena gestión de la recopilación de datos requiere que los administradores trabajen en estrecha colaboración con los encuestadores, especialmente durante los primeros días de la recolección. Es imperativo que los administradores comprueben minuciosamente la calidad de la recopilación de datos. El director de la investigación debe validar las cajetillas y las fotografías tomadas. De esta manera, los errores pueden abordarse desde el principio y corregirse en consecuencia sin desperdiciar recursos. Por ejemplo, en el estudio brasileño se puso de manifiesto durante los primeros días que los encuestadores no podían entrar en los edificios de los hogares más ricos para realizar la encuesta de análisis de cajetillas. Las altas medidas de seguridad impidieron que los encuestadores pudieran entrar. El gerente de la investigación decidió que los encuestadores deberían esperar a los residentes fuera del edificio. Con esta modificación, la recopilación de datos continuó con poca interrupción, y el número adecuado de los hogares más ricos se siguió incluyendo en la encuesta (Carvalho Figueiredo, próxima publicación).

Como se examinó en la Sección 3.2.3, en algunos estudios sobre el comercio ilícito se utilizan simultáneamente el método de análisis de cajetillas de fumadores y el de recolección de cajetillas desechadas, lo que permite realizar una validación cruzada de los resultados. Cuando se utiliza la probabilidad proporcional al tamaño (PPT) para muestrear áreas (UPM), ambos métodos pueden llevarse a cabo simultáneamente. El encuestador puede recoger las cajetillas desechadas y realizar encuestas a lo largo de la ruta. Sin embargo, como la recolección de cajetillas desechadas requiere menos habilidad, la aplicación de esta estrategia solo es eficiente cuando el encuestador hace muchas caminatas entre los hogares. De lo contrario, puede ser más eficiente tener dos grupos de encuestadores, uno centrado en la recolección de cajetillas desechadas y el otro en la realización de las encuestas. Por ejemplo, en un estudio sobre el comercio ilícito realizado en Varsovia (Polonia) se contrató a una empresa de estudio de mercados para que realizara entrevistas con fumadores, mientras que se trabajó con estudiantes universitarios para recolectar cajetillas desechadas (Stoklosa y Ross, 2014). Cuando los dos métodos se llevan a cabo por separado, deben aplicarse en un plazo similar para permitir la validación cruzada.

La entrada de datos es la etapa final del proceso de recopilación de datos. Los administradores de datos supervisan la creación de un conjunto de datos factibles. En este conjunto de datos, cada cajetilla representa una entrada. Hay varias maneras de imputar la información. Estos conjuntos de datos pueden crearse en un sistema de gestión de bases de datos en un software estadístico (por ejemplo, Stata), o en un programa de hoja de cálculo (por ejemplo, Microsoft Excel). El método de recolección de cajetillas desechadas requiere que cada cajetilla sea evaluada individualmente y que la información se ingrese en la base de datos. La encuesta de análisis de cajetillas de fumadores requiere que la información se transcriba en la base de datos. Las características de los fumadores, incluida la intensidad con la que fuman, también se incluirán en cada cajetilla. Para ambos métodos, es necesario añadir a cada entrada el tiempo y el lugar (por ejemplo, la UPM) de la recopilación de datos. Si se dispone de información adicional, como las coordenadas y las imágenes del GPS, también se debe incluir esta información. Una buena práctica consiste en incluir una referencia a la fotografía de cada cajetilla en la entrada de datos para facilitar su identificación posterior.

No es necesario introducir todas las características de la cajetilla en la base de datos. En el Gráfico 2 de la Sección 3.2.2 se presenta un enfoque secuencial para identificar las cajetillas ilícitas. Concretamente, en un ejemplo presentado en ese gráfico, se examina primero una marca de cigarrillos. Las cajetillas con marcas que no están registradas para la venta legal se codifican como ilícitas. Solo las cajetillas con marcas que están oficialmente permitidas en el mercado se someten a un examen detallado. En este enfoque, la información sobre las marcas se utiliza para eliminar rápidamente todas las marcas que son ilícitas. El enfoque secuencial les ahorra tiempo a los codificadores, ya que no necesitan examinar otras características de las cajetillas con marcas ilícitas.

Tengan cuidado. Los errores se colarán en cada etapa del proceso. Es especialmente vital comprobar si hay errores en la base de datos. Para comprobar el método de recolección de cajetillas desechadas es aconsejable que dos codificadores distintos ingresen la información de un subconjunto de las cajetillas recolectadas. Si por alguna razón los datos ingresados no coinciden, entonces los investigadores pueden investigar más a fondo. De lo contrario, puede ser difícil encontrar errores. Con la encuesta de análisis de cajetillas, se puede cotejar un subconjunto de la muestra comparando la información ingresada en la base de datos con las fotografías de las cajetillas que tomaron los encuestadores.

3.4.4 Identificación de cajetillas falsificadas

Es posible que algunas cajetillas de cigarrillos parezcan legales, pero en realidad se producen ilegalmente. Estos tipos de cigarrillos caen en la categoría de cajetillas falsificadas. Las cajetillas falsificadas son relativamente raras. Los informes de la industria tabacalera afirman que solo el 0,5% de los cigarrillos consumidos en todo el mundo son falsificados (Philip Morris International, 2019). No obstante, varios países se enfrentan a un problema de falsificación de cigarrillos.

Cuando la falsificación es un problema, la irradiación ultravioleta (UV) y la microscopía óptica pueden identificar cigarrillos falsificados (Kurti *et al.*, 2017). El método fue desarrollado por Kurti. El equipo necesario incluye un microscopio óptico con una cámara USB, una lámpara UV, un gabinete de visión UV y, opcionalmente, un sensor UV USB. El equipo es relativamente barato y la aplicación es sencilla. El método Kurti utiliza una prueba de diagnóstico realizada en cajetillas falsificadas proporcionadas por los organismos de orden público.

Los investigadores deben obtener cajetillas falsificadas confirmadas (por ejemplo, de los organismos de orden público) que puedan luego compararse con las cajetillas legales confirmadas (por ejemplo, obtenidas de tiendas confiables) para calibrar una prueba con una especificidad y sensibilidad óptimas. Con la calibración completa, se puede implementar la segunda fase del proyecto. En esta fase, las cajetillas se recogen a través de la recolección de cajetillas desechadas o a través de la encuesta de análisis de cajetillas de fumadores. Esas cajetillas están ordenadas por su legalidad. Una cajetilla que se considere legal después del análisis inicial será sometida a nuevas pruebas para determinar definitivamente que no sea una falsificación. La prueba de falsificación se utiliza para establecer la legalidad de la cajetilla.

Las características de las cajetillas de cigarrillos falsificados pueden diferir entre los países. El método Kurti fue diseñado específicamente para identificar los cigarrillos Marlboro falsificados. Este método también ha demostrado ser adecuado para verificar otras marcas de cigarrillos (Kurti *et ál.*, 2017). El método Kurti sigue siendo apropiado para la identificación de cajetillas falsificadas en una multitud de ámbitos. En el Gráfico 9 a continuación, se describen detalladamente los lineamientos del método Kurti.

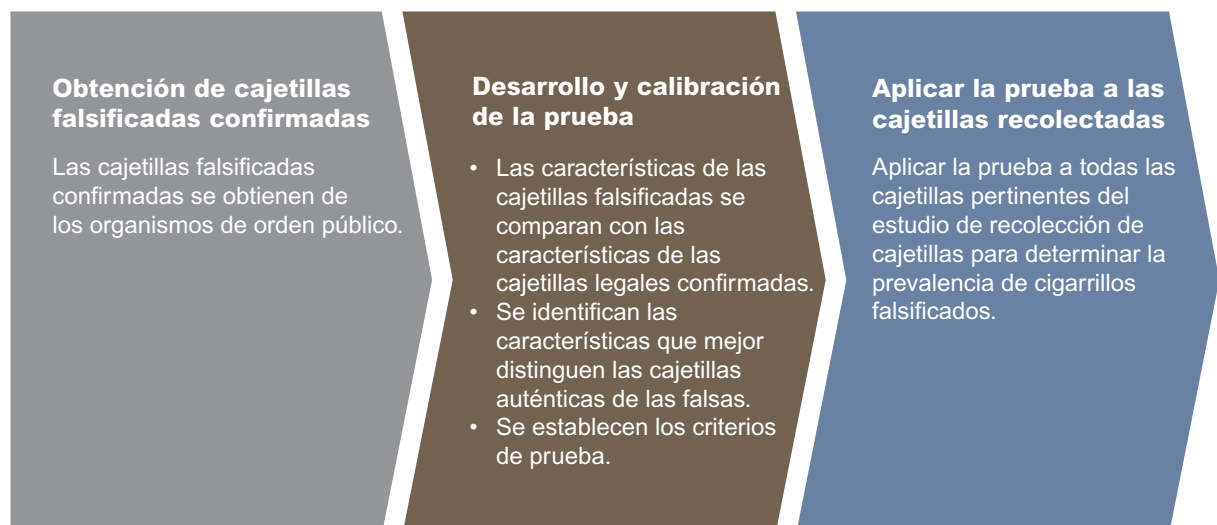
3.4.4.1 Creación y calibración de la prueba

La primera etapa del método Kurti requiere la colaboración con los organismos de orden público. Los investigadores deben identificar las marcas de cigarrillos falsificados más frecuentes en su área de estudio. Los organismos de orden público tendrían acceso a esta información, así como a muestras de cajetillas falsificadas. Las cajetillas de cigarrillos falsificadas deben ser las más recientes. La mayoría de las cajetillas falsificadas se habrán incautado mediante redadas en instalaciones de fabricación ilegal. Lo mejor sería que las muestras de las cajetillas falsificadas de cada marca provengan de incautaciones diferentes. Estas cajetillas obtenidas de los organismos de orden público constituyen la muestra de cajetillas falsificadas confirmadas independientemente.

Los investigadores también deben utilizar un grupo de control para la comparación. Estas cajetillas legales también se deben verificar como legales. Cada marca falsificada presente en la muestra debe tener una contraparte legal. Las cajetillas legales se pueden obtener a través de cadenas de venta al por menor de confianza o directamente de los fabricantes de tabaco que operan legalmente.

Una vez que se hayan obtenido todas las cajetillas necesarias, es preciso reunir datos sobre las características de las cajetillas de ambas muestras. La radiación ultravioleta y la microscopía óptica se utilizan para examinar cada cajetilla. Siguiendo la descripción del método Kurti (Kurti *et ál.*, 2017), el paquete exterior, el marco interior y, si la cajetilla está envuelta en papel de aluminio, el papel de aluminio interior, deben ser expuestos a una luz UV de

Gráfico 9 Desarrollo de la prueba de cajetillas falsificadas



onda larga. Los investigadores deben reunir datos sobre la presencia de agentes blanqueadores ópticos y fluorescencia para ambos grupos. El sensor UV puede medir objetivamente la presencia de ambas características. En segundo lugar, los investigadores deben buscar signos de impresión de mala calidad en varios lugares predefinidos de cada cajetilla falsificada utilizando el microscopio óptico. Estas áreas se comparan con la impresión de las cajetillas genuinas. Cada marca tendrá elementos únicos en el diseño de la impresión que distingan las cajetillas falsificadas de las legales.

Con la verificación microscópica y los datos de irradiación UV, los investigadores pueden identificar las características específicas de la marca que mejor distinguen las cajetillas genuinas de las falsificadas. La prueba ideal tendrá la mayor especificidad (es decir, la capacidad de detectar si una cajetilla es genuina, o verdaderos negativos) así como la mayor sensibilidad (es decir, la capacidad de detectar si la cajetilla es falsa, o verdaderos positivos). Los investigadores deberían utilizar el análisis de la Curva de las Características de Operación del Receptor (ROC, por sus siglas en inglés). Esto ayudará a identificar las características óptimas de cada marca.

Es importante señalar que esta prueba no es 100% exacta para distinguir entre las cajetillas legales y las falsificadas. La especificidad y sensibilidad de la prueba proporcionará a la investigación un margen de error. Se ha comprobado que el método Kurti tiene una gran precisión entre marcas, ya que la precisión de la prueba es similar entre estas. La prueba clasificó correctamente el 100% de las cajetillas de Newport y Marlboro Gold usando solo dos características de la cajetilla (sensibilidad y especificidad del 100%). La prueba identificó correctamente el 96% de la muestra regular de Marlboro utilizando cuatro características (sensibilidad del 94,12% y especificidad del 100%) (Kurti *et ál.*, 2017). Es esencial que los clasificadores de cajetillas estén bien capacitados para que puedan recopilar eficazmente la información necesaria de las cajetillas de cigarrillos mediante esta prueba.

3.4.4.2 Determinar el alcance de la falsificación de cigarrillos mediante la prueba

Una vez elaborada la prueba de falsificación, se puede aplicar a una gran muestra de cajetillas obtenidas de una colección de cajetillas desechadas o de una encuesta de análisis de cajetillas de fumadores en la que estos proporcionan sus cajetillas. La prueba se aplicará a las cajetillas consideradas legales. Estas cajetillas se identifican como legales porque presentan advertencias sanitarias y/o sellos fiscales apropiados claramente visibles y fácilmente reconocibles. La prueba ayudará a determinar si una cajetilla clasificada como legal es de hecho genuina.

La prueba solo podrá distinguir entre las marcas que han sido analizadas. Por ejemplo, si solo se utilizaron los cigarrillos de Pall Mall para desarrollar la prueba, esta solo puede comprobar si las cajetillas de cigarrillos Pall Mall son legales o falsas. Una buena práctica es hacer que el 10% de la submuestra legal sea evaluada independientemente por otra calificadora. La confiabilidad y la precisión de las pruebas pueden confirmarse mediante esta técnica.

3.4.5 Uso de estudios de análisis de cajetillas para otros productos de tabaco

Aunque los métodos descritos en el presente capítulo se refieren únicamente a los cigarrillos, tanto la encuesta sobre el análisis de cajetillas de cigarrillos de fumadores como la recolección de cajetillas desechadas se pueden utilizar para medir la prevalencia del comercio ilícito de otros productos de tabaco. En los lugares donde el empaque de los demás productos está regulado para exigir, por ejemplo, advertencias sanitarias o sellos fiscales, los investigadores pueden reunir datos y evaluar los empaques utilizando las técnicas descritas en este capítulo. Sin embargo, hasta ahora no se ha realizado ningún estudio a gran escala sobre el comercio ilícito de productos de tabaco distintos a los cigarrillos.

3.5 Análisis de datos

La característica de identificación más importante de cada cajetilla es si cumple o no con las normas. Esto se puede representar fácilmente con una variable dicotómica (cero para las que cumplen, uno para las que no cumplen). Asignar la variable dicotómica correcta puede parecer trivial, pero no siempre es sencillo saber qué

cajetillas cumplen y cuáles no. Ocasionalmente, se deben tener en cuenta múltiples características de la cajetilla para determinar la legalidad de esta. En la Sección 3.2.2 se examina a fondo la forma de utilizar las características de las cajetillas para distinguir las legales (que pagan impuestos) de las ilegales (que no pagan impuestos).

Las encuestas de análisis de cajetillas de fumadores en países con una alta proporción de fumadores de cigarrillos individuales deben aplicar una condición de cumplimiento diferente, ya que los fumadores pueden no tener una cajetilla disponible. En este caso, la información sobre el precio de los últimos cigarrillos comprados, la marca o el nombre del cigarrillo y otras características se pueden utilizar para calcular la situación de cumplimiento de estos fumadores. El estudio sobre el comercio ilícito de cigarrillos en México, por ejemplo, determinó que el nombre de la marca de cigarrillos era un indicador suficiente de la legalidad de los cigarrillos individuales (Sáenz de Miera Juárez, 2020). Cuando la información sobre marcas y precios es insuficiente para determinar el estado de cumplimiento, se puede imputar el estado a los fumadores que muestran sus cajetillas.

Un modelo econométrico es un enfoque más riguroso para asignar la condición de cumplimiento a los fumadores que no muestran sus cajetillas. En este caso, se implementa un modelo predictivo para estimar los valores que faltan. Las características de los consumidores y la información sobre los cigarrillos son necesarias para estimar esos valores que faltan. Las variables que mejor predicen el cumplimiento suelen ser específicas de cada país y pueden variar. El Anexo 6.7 proporciona un código de Stata que predice el estado de cumplimiento para los fumadores que no muestran sus cajetillas.

El siguiente paso es estimar la prevalencia de las cajetillas que no cumplen con las normas en una región geográfica determinada. Como demostró Merriman, en las recolecciones de cajetillas desechadas, el promedio matemático simple de las probabilidades de cumplimiento de cada una de las UPM es un estimador no sesgado de la verdadera probabilidad media regional de cumplimiento (Merriman, 2010). La prevalencia de cajetillas que no cumplen en la muestra es una buena representación de la prevalencia de cajetillas que no cumplen en la región. El cálculo de la prevalencia de las cajetillas que no cumplen con los requisitos de la encuesta de análisis de cajetillas de fumadores es solo un poco más complicado. Primero, los investigadores deben ponderar el estado de cumplimiento de la cajetilla de cigarrillos por la intensidad del consumo de tabaco del encuestado. Esto asegura que la ponderación es proporcional al número de cigarrillos que se fuman por día. En segundo lugar, si el muestreo estratificado dio lugar a que algunos estratos fueran relativamente más pequeños, se pueden aplicar ponderaciones para ajustar las submuestras más pequeñas. La prevalencia del incumplimiento puede entonces computarse con precisión.

Además de estimar la prevalencia de los cigarrillos ilícitos, los investigadores tal vez deseen realizar análisis adicionales de los datos. Las pruebas de proporciones (comando `<prtest>` en Stata) se pueden utilizar para determinar si la probabilidad de fumar cigarrillos que no cumplen con las normas difiere entre el género, la edad o los grupos socioeconómicos. La prueba de Pearson χ^2 (comando `<tabulate, chi2>` en Stata) se puede utilizar para determinar si los precios son diferentes entre los cigarrillos que cumplen y los que no cumplen. Por último, el análisis multivariado se puede utilizar para estudiar relaciones más complejas en los datos. Por ejemplo, Merriman utiliza un modelo probit (comando `<probit>` en Stata) para analizar si la probabilidad de cumplimiento fiscal en una determinada UPM depende de la proximidad de la UPM a una jurisdicción con menores impuestos. Estos cálculos controlan los ingresos del hogar, la raza y otras características socioeconómicas de los habitantes de la UPM (Merriman, 2010).

3.6 Conclusión

En este capítulo se presentan métodos de análisis de cajetillas para estimar la prevalencia del comercio ilícito de productos de tabaco. Se centra en dos métodos que se aplican comúnmente en los estudios sobre el comercio ilícito: la encuesta de análisis de cajetillas de fumadores y la recolección de cajetillas desechadas. El contenido de este capítulo debe dotar a los investigadores de conocimientos sobre todos los aspectos del estudio, desde la planeación de la investigación y el diseño de la muestra, pasando por la recopilación de datos, hasta el análisis de estos.

El análisis de cajetillas mediante la recopilación de datos primarios es la forma más confiable de medir el comercio ilícito de tabaco, pero también es costoso y lleva mucho tiempo. El siguiente capítulo se centra en el análisis de los datos secundarios para estimar el comercio ilícito de tabaco.

4

Modelo de brecha

4.1 Introducción

Un modelo de brecha estima la evolución de la tendencia del comercio ilícito de tabaco. Este modelo estima la diferencia entre el mercado total de cigarrillos (combinación de impuestos pagados e impuestos no pagados) según lo medido por las encuestas y los volúmenes de cigarrillos que reportaron haber pagado impuestos.

La mayoría de los países reúnen información de encuestas de la prevalencia de tabaquismo (el porcentaje de la población que fuma en un período determinado) y la intensidad del tabaquismo (el número de cigarrillos que fuma cada fumador en un período determinado). Los países también deben disponer de cifras precisas, o al menos de estimaciones adecuadas, del tamaño total de la población. La brecha puede medirse utilizando estas estimaciones. El número de cigarrillos consumidos puede estimarse combinando la prevalencia de tabaquismo, la intensidad de este y las estimaciones de la población. Como suele ocurrir con las técnicas de estimación, los resultados solo sirven como una aproximación del consumo total de cigarrillos; el error de medición presente en la encuesta llevará a la aproximación. Además del error de medición de la encuesta, la gente puede no informar con precisión sus hábitos tabáquicos. Esto introduce un sesgo potencialmente no aleatorio.

Si bien la aplicación del análisis de brechas no es teóricamente difícil, la aplicación del análisis de brechas puede ser a menudo difícil debido a las limitaciones de los datos y el análisis. Además, el análisis y la interpretación pueden ser arduos. La comprensión de los puntos fuertes y débiles de este enfoque es importante durante la etapa de análisis e interpretación. Este capítulo abarca tanto los conceptos teóricos como la interpretación, y también aborda los matices más finos relacionados con la interpretación y la presentación de los resultados.

En primer lugar, se explican los fundamentos conceptuales del modelo de brecha, basado en las identidades de consumo. En segundo lugar, se explican conceptualmente los puntos fuertes y las limitaciones del modelo, y se entregan casos empíricos. En tercer lugar, se examinan las fuentes de datos necesarias para utilizar este método y se dan ejemplos. En cuarto lugar, se explican con ejemplos concretos varios problemas que los datos pueden presentar y que se deben examinar. En quinto lugar, se dan ejemplos del código de Stata para aplicar la metodología, junto con diferentes formas de presentar los resultados. Por último, se presentan varios casos prácticos en los que se ha aplicado este método.

4.2 Marco conceptual del modelo de brecha

Sean Q^T el número total de cigarrillos consumidos en un determinado período en un país o jurisdicción relevante (con una política fiscal común sobre el tabaco), Q^L la cantidad total de cigarrillos que pagan impuestos al tabaco en el mismo período/país, y Q^E los cigarrillos que se producen y pagan impuestos a nivel nacional pero que se introducen de contrabando fuera del país:

$$Q^T = (Q^L - Q^E) + Q^I$$

donde Q^I es la cantidad de cigarrillos sin pagar impuestos consumidos en ese período/país. Los cigarrillos que no pagan impuestos no son necesariamente cigarrillos ilícitos, ya que pueden incluir, por ejemplo, consumo personal libre de impuestos. Se supone que la proporción de cigarrillos lícitos que no pagan impuestos del Q^I se mantiene constante a lo largo del tiempo. Las fluctuaciones del Q^I se deben únicamente a las fluctuaciones del consumo ilícito de cigarrillos. Q^I es la variable de interés y es la brecha que será estimada y comparada a lo largo del tiempo:

$$Q^I = Q^T - (Q^L - Q^E) \quad (1)$$

A su vez, Q^T se puede definir como el mercado de consumo interno total (Q^D) y es igual a:

$$Q^T \equiv Q^D = \sum_{i=1}^S N_i$$

donde S es el número total de fumadores en el período relevante en un país/jurisdicción y N es el número de cigarrillos consumidos por el fumador i en dicho período/país. Al trabajar con datos agregados:

$$Q^D = P^T \times M^{Prev} \times \bar{N}$$

P^T es la población total de un país en un determinado período, M^{Prev} es la proporción de fumadores en la población de ese período/país, y \bar{N} es el promedio de cigarrillos fumados en ese período/país. M^{Prev} y \bar{N} se tienen que medir en el mismo período de tiempo, como por ejemplo, por mes. Entonces, M^{Prev} sería la prevalencia mensual de tabaquismo y \bar{N} el número promedio de cigarrillos fumados mensualmente por esa población. Si la intensidad promedio se mide semanalmente, pero la prevalencia se mide mensualmente, entonces debe suponerse que la intensidad semanal es constante dentro del mes. El promedio de cigarrillos que se fuman al mes es 4,43 veces el promedio semanal.

La identidad expresada en (1) no se observa en la práctica debido a que los fumadores subregistran sus hábitos tabáquicos (en N). A menudo el resultado es que Q^L es mayor que Q^D , un resultado sin sentido. Por consiguiente, la relación en (1) debe evaluarse a lo largo del tiempo para evaluar si el subregistro (y la proporción de cigarrillos libres de impuestos consumidos legalmente) es constante. La ecuación (1) debe transformarse a:

$$\frac{\Delta Q^I}{Q_t^I} = \frac{\Delta(Q^D)}{(Q_t^D)} S - \frac{\Delta Q^L}{Q_t^L} (s - 1) \quad (2)$$

donde Δ es la diferencia de la cantidad relevante en los períodos $t+1$ y t (año base). Cada relación representa el porcentaje de aumento en el tiempo t . La variable s representa la relación entre la cantidad total del mercado de consumo interno y el mercado ilícito en el tiempo t (Q_t^D / Q_t^I). Por tanto, s es una constante mayor que 1 y, naturalmente, desconocida. La ecuación (2) resuelve s y presenta la evolución del mercado ilícito.

Si el aumento porcentual del mercado interno total (Q^D) es mayor (o menor) que el aumento porcentual en Q^L , indicaría que Q^I ha aumentado (o disminuido) proporcionalmente con respecto al año base. Suponiendo que el subregistro es constante en el tiempo, tal aumento (o disminución) sería indicativo de un aumento (o disminución) en la cuota del mercado del comercio ilícito (definido como comercio ilícito y como porcentaje del mercado total).

Si todo sigue igual, puede producirse un aumento relativo de la cuota del mercado del comercio ilícito en un mercado de consumo interno que aumenta con el tiempo. Por lo tanto, habría un aumento absoluto de la

cuota del mercado ilícito. Por otra parte, la disminución de la cuota del mercado del comercio ilícito y del mercado de consumo interno implicaría una reducción del volumen de cigarrillos ilícitos que se venden. Por último, si el mercado de consumo interno total está aumentando (disminuyendo) y la cuota del mercado del comercio ilícito está disminuyendo (aumentando), no se puede calcular la evolución del volumen absoluto vendido en el mercado ilícito.

Solo se requiere una estimación del tamaño del mercado ilícito (en volumen) para extrapolar el tamaño del mercado ilícito en otros puntos del tiempo. Las condiciones de estimación requieren necesariamente que el subregistro sea constante a lo largo del tiempo. En un estudio realizado en Brasil, los investigadores utilizaron la estimación del volumen total de cigarrillos vendidos en el mercado ilícito en 2013 para deducir el volumen correspondiente a 2012, 2014, 2015 y 2016 (Szklo *et ál.*, 2017).

Cuando no se dispone del volumen real del comercio ilícito, medido con datos primarios (encuesta de análisis de cajetillas), se puede suponer un cierto porcentaje fijo de subregistro. El porcentaje fijo puede añadirse a la tendencia estimada mediante el método de brecha (Blecher, 2010). Los indicadores indirectos deben sugerir que la proporción de subregistro se encuentra dentro de los límites aproximados y razonables, o los resultados no serán confiables. Los análisis de sensibilidad proporcionarían solidez en el porcentaje de subregistro.

4.3 Fortalezas, debilidades y limitaciones

La mayor fortaleza del método de brecha radica en lo económico de los datos y recursos. En cuanto a los datos, solo se necesitan dos conjuntos de datos. El primer conjunto de datos necesario es la cantidad de cigarrillos con impuestos pagados. Estos datos suelen estar disponibles en las autoridades fiscales nacionales y/o en las oficinas nacionales de aduanas de los países en que los cigarrillos se importan en su totalidad.

El segundo conjunto de datos que se necesita es el de los patrones de consumo de cigarrillos. Estos datos suelen obtenerse mediante encuestas a los consumidores. Los gobiernos recopilan estos datos de forma periódica. La información requerida es la prevalencia de tabaquismo durante un período determinado (por ejemplo, el mes anterior al estudio) y la intensidad promedio del tabaquismo durante el mismo período o un período más corto (por ejemplo, la semana anterior al estudio). Es relativamente fácil acceder a los estudios que contienen las dos variables. A menudo, los gobiernos archivan sus datos. Esto hace que sea posible la estimación de las tendencias históricas, aunque también se necesitarán datos sobre las ventas con los impuestos pagados para estimar las tendencias históricas.

Por lo que respecta a los recursos, el método de brecha no requiere la recopilación de datos primarios y se puede realizar utilizando únicamente datos secundarios. Tampoco se requieren programas estadísticos especiales ni amplias capacidades econométricas. El método de brecha solo requiere un software de hoja de cálculo simple y conocimientos estadísticos básicos.

El método de brecha tiene dos debilidades principales. En primer lugar, el método no proporcionará una estimación precisa ni del volumen ni de la proporción. El principal problema de las estimaciones es la falta de información. Se supone que el subregistro es constante a lo largo del tiempo y es en sí mismo una estimación. Ambas suposiciones son muy fuertes. Es posible que el tamaño del subregistro sea significativo. Por ejemplo, en un estudio realizado en Argentina, las cantidades de cigarrillos reportadas en las encuestas entre la población general (mayores de 18 años) son entre un 25% y un 35% más bajas que las ventas registradas con impuestos pagados (Paraje, 2018). En el caso de Brasil, pueden ser hasta un 17% más bajos (Paraje, 2018).

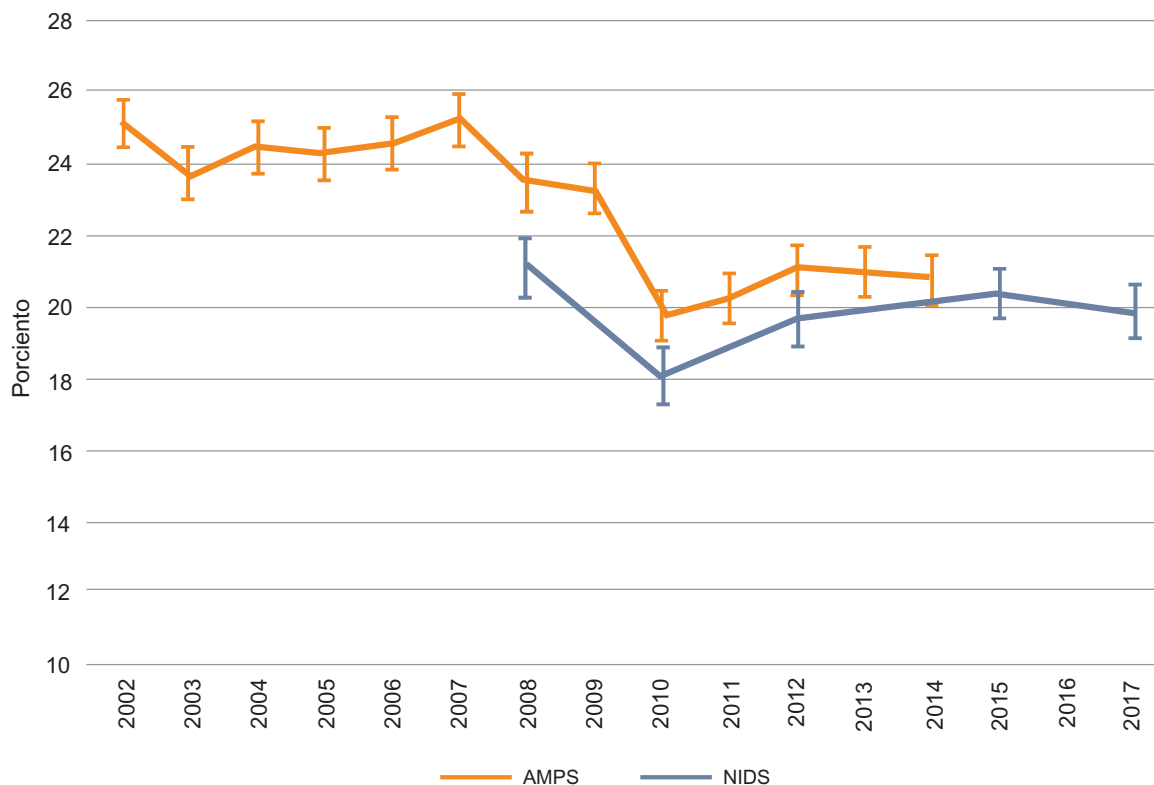
A menudo es imposible calibrar el grado de subregistro. Por lo tanto, el método de brecha no proporciona una cifra exacta o un porcentaje del tamaño del mercado ilícito. Si se puede suponer que el subregistro del consumo de cigarrillos es una constante o una proporción del consumo registrado, entonces el análisis de brecha puede estimar las tendencias del consumo de cigarrillos en los casos en que no se han pagado impuestos. Si no se puede hacer esa suposición de forma razonable, no se debe utilizar el análisis de brecha.

En segundo lugar, el análisis de brecha solo se debe utilizar si la proporción fija de subregistro (y la proporción constante para las compras libres de impuestos, etc.) es realista. Dado que el subregistro es una incógnita, *a priori*, la suposición debe estar respaldada por afirmaciones razonables. La suposición es más creíble cuando se dispone de mediciones coherentes y regulares del consumo de cigarrillos de la población (por ejemplo, encuestas a los consumidores). La coherencia dentro del marco de muestreo de la encuesta y los métodos de recolección es necesaria para fines comparativos. La suposición de que los cigarrillos de contrabando salen del país también debe ser constante. Al igual que con los cigarrillos de contrabando en el país, esta cantidad es raramente conocida.

El método de brecha proporciona resultados más confiables cuando se utilizan estudios consistentes, o estudios con metodologías similares, durante algún periodo de tiempo. No es aconsejable utilizar encuestas con diferentes marcos de muestreo, diseños de encuestas, variaciones estacionales, definiciones variables del consumo de tabaco, productos de tabaco o edades. Es probable que la utilización de encuestas que varíen en función de esos criterios dé lugar a inconsistencias o a dificultades para comparar las estimaciones a lo largo del tiempo. Por ejemplo, la Encuesta sobre toda la Media y Productos de Sudáfrica y el Estudio sobre la Dinámica de los Ingresos Nacionales estiman la prevalencia de tabaquismo en años traslapados. El Gráfico 10 muestra los resultados de la prevalencia de tabaquismo traslapados en tres años (2008, 2010 y 2012). El Estudio de la Dinámica de los Ingresos Nacionales estima sistemáticamente una menor prevalencia de tabaquismo en los años traslapados.

El método debería definir lo que se está midiendo. El consumo total estimado (a menudo subestimado) se suele estimar a partir de encuestas sobre el consumo de tabaco o de drogas con un módulo para tabaco. Las encuestas se realizan a partir de la población general o de grupos específicos (por ejemplo, adultos). Es raro que las encuestas tengan información sobre el origen de los cigarrillos. Esta información sería útil porque proporcionaría

Gráfico 10 Prevalencia de tabaquismo en Sudáfrica, 2002-2017



Origen: Vellios *et al.*, (2019).

datos sobre si los fumadores consumen cigarrillos lícitos o ilícitos. Es posible que los propios consumidores desconozcan la situación de los cigarrillos que consumen, como es el caso de los fumadores de cigarrillos individuales. La cantidad total de cigarrillos consumidos proporcionada incluye los cigarrillos lícitos e ilícitos.

Cuando se dispone de datos sobre el consumo de cigarrillos solo para un subconjunto no representativo de la población (por ejemplo, jóvenes, mujeres/hombres, habitantes de una región determinada, etc.), los datos sobre los impuestos pagados deben corresponder a ese grupo (por ejemplo, los cigarrillos pagados por la población adulta). Si no se dispone (con bastante frecuencia) de información sobre los cigarrillos con impuestos pagados para ese subconjunto de la población, no se puede realizar un análisis de brecha.

4.4 Fuentes de datos

4.4.1 *Estimación de las ventas con impuestos pagados utilizando datos oficiales*

Muchos gobiernos publican datos sobre el volumen de ventas de cigarrillos y/o los ingresos por concepto de impuestos a los cigarrillos. Los datos sobre el volumen pueden recopilarse y presentarse directa o indirectamente mediante la medición de marcas fiscales como los sellos fiscales.

Muchos otros países no tienen datos sobre el volumen de ventas y solo publican datos sobre los ingresos. Los datos sobre el volumen pueden ser fácilmente imputados cuando el sistema se basa en impuestos específicos. Por ejemplo, si se reciben \$1.000 millones de dólares en ingresos en un año y se aplica un impuesto específico de \$0,50 dólares por cajetilla, entonces el volumen de cigarrillos es de 2.000 millones de cajetillas al año ($1.000.000.000 / 0,50$). Cuando se utilizan sistemas fiscales específicos por niveles, los datos de ventas imputados solo se pueden derivar si los ingresos se especifican por niveles. Como alternativa, se pueden utilizar los datos sobre la cuota del mercado externo, si está disponible.

Los sistemas fiscales *ad valorem* presentan retos, ya que la base impositiva puede ser difícil de comprender o estimar. Además, los ingresos no se pueden utilizar para determinar el volumen. En cambio, es posible utilizar la base impositiva promedio sobre la que se aplican los impuestos. A partir de la base impositiva, los impuestos recaudados se pueden utilizar para extrapolar el número de cigarrillos. Suponiendo que los ingresos *ad valorem* son de \$1.000 millones de dólares, el precio promedio de venta al público es de \$1 dólar y la cuota promedio de impuestos al tabaco es del 50%, lo que significaría que la base impositiva es de \$0,50 dólares. La relación entre la base impositiva y los ingresos totales podría dar una estimación del total de cigarrillos legales consumidos. Con este enfoque, es necesario hacer una suposición especialmente fuerte. La composición de la base impositiva, en términos de marcas, permanece fija a lo largo del tiempo. Esto puede ser potencialmente una suposición poco realista.

Los cambios en los impuestos durante el período de estudio requerirán ajustes. Los datos del gobierno especificarán la cobertura temporal. Si no se dispone de datos anuales, los investigadores deben tener cuidado al imputar los datos anuales debido a las fluctuaciones estacionales, las políticas de almacenamiento específicas de las compañías tabacaleras, etc.

4.4.2 *Estimación de las ventas con impuestos pagados utilizando fuentes externas*

En muchos países, los datos agregados sobre las ventas de cigarrillos se pueden obtener de empresas de investigación de mercados del sector privado, entre ellas Euromonitor International o GlobalData. Estos datos anuales suelen presentarse como una serie temporal histórica. Sin embargo, estos datos no están exentos a desafíos. Abarcan un número limitado de países, aunque por lo general se dispone de datos sobre países grandes. Además, los conjuntos de datos varían. Las revisiones de los datos actuales e históricos sugieren utilizar los conjuntos de datos más actualizados. Las fuentes de datos también pueden variar y pueden proceder de datos del gobierno o de la industria. A menudo los datos pueden ajustarse a los años calendario, más que a los años fiscales, o al marco temporal de los datos de la encuesta.

Otras instituciones gubernamentales o internacionales también pueden ser útiles para encontrar datos de consumo externos, entre ellas el Departamento de Agricultura de los Estados Unidos y la Organización de las Naciones Unidas para la Agricultura y la Alimentación. Sin embargo, al utilizar estos tipos de datos se debe tener en cuenta la fuente de estos. Los datos de consumo se pueden imputar a partir de los datos de producción y comercio de hojas de tabaco o de los datos de comercio de tabaco manufacturado. Con estas fuentes, al igual que con la base de datos Comtrade de las Naciones Unidas, el consumo de cigarrillos que pagan impuestos en el ámbito nacional se estima utilizando la relación de consumo aparente, definida como:

$$C_t = P_t - X_t + M_t$$

donde C_t es el consumo (que paga impuestos) en el tiempo t , P_t es tabaco manufacturado en el tiempo t , X_t es exportación en el tiempo t , y M_t es importación (que paga impuestos) en el tiempo t . A los datos hay que aplicarles una tasa de conversión de cigarrillos individuales invariante en el tiempo, típicamente medida en toneladas. Los investigadores podrían asumir que un cigarrillo individual pesa un gramo. Una tonelada de cigarrillos significa que hay 1.000.000 de cigarrillos. Sin embargo, esta suposición puede no ser realista durante largos períodos de tiempo, ya que la cantidad de tabaco en los cigarrillos manufacturados ha disminuido con el tiempo.

El consumo aparente debe estimarse en función de las cantidades de cigarrillos fabricados. No es apropiado utilizar otros productos de tabaco como hoja, tabaco crudo, etc. Los códigos del Sistema Armonizado de Denominación y Codificación de Mercancías (HS, por sus siglas en inglés), para estos productos son: (240210) Puros, minipuros y purillos que contengan tabaco, incluido el peso de cada banda, envoltorio o accesorio; y (240220) Cigarrillos que contengan tabaco, etc.

4.4.3 Encuestas sobre la prevalencia de tabaquismo

Las encuestas sobre el consumo de tabaco o sobre el consumo de drogas con un módulo para tabaco se utilizan para obtener datos sobre la prevalencia de tabaquismo (en un determinado período, por lo general el mes anterior a la encuesta) y sobre la intensidad del tabaquismo. En América del Sur, por ejemplo, en las encuestas se pregunta sobre el consumo de tabaco en el mes anterior a la encuesta y el promedio de consumo diario en el mismo período. En al menos cinco países, estas encuestas se han realizado por lo menos dos veces, con preguntas comparables, lo que permite utilizarlas para hacer comparaciones a lo largo del tiempo (Paraje, 2018).

Una cuestión que pudiera significar desafíos es la población cubierta por una encuesta. Por ejemplo, el Estudio Nacional de Drogas en Población General de Chile, que se realiza cada dos años, encuesta a personas de 12 a 65 años (Observatorio Chileno de Drogas, 2017). El Estudio Nacional de Drogas en Población Escolar de Chile, que también se realiza cada dos años, reúne información sobre el consumo de tabaco en niños que suelen tener entre 12 y 18 años (Observatorio Chileno de Drogas, 2015). Naturalmente, estas dos encuestas no se pueden utilizar conjuntamente, ya que el marco de muestreo es diferente. Sería aconsejable utilizar la encuesta con la cobertura más amplia de la población de fumadores. En este caso, sería preferible la Encuesta de Población en General. Sin embargo, esta encuesta no incluye a los fumadores de 65 años o más. En general, se supone que la prevalencia e intensidad del consumo del grupo excluido es constante a lo largo del tiempo.

Por otra parte, en la Encuesta Nacional de Factores de Riesgo de Argentina, el marco de muestreo incluye a la población de 18 años y más (Instituto Nacional de Estadísticas y Censos, 2015). Se excluye a la población menor de 18 años. Sin embargo, la edad promedio de inicio en el tabaquismo en Argentina se acerca a los 16 años (Guindon et ál., 2017). Hay que suponer que los jóvenes de 16 a 18 años tienen patrones de consumo idénticos e invariantes en el tiempo en cuanto a la prevalencia e intensidad de tabaquismo. Esta suposición puede parecer inocua para la población mayor de 65 años (ya que son fumadores de largo plazo, potencialmente menos influenciados por las políticas de control del tabaco), pero podría ser más arriesgada para la población menor de 18 años. Para reducir los riesgos, las encuestas se deben realizar en momentos cercanos, lo que hace menos probable que cualquier cambio de política surta efecto pleno y afecte la prevalencia o la intensidad.

Estas encuestas se pueden consultar con frecuencia en los sitios web de las Oficinas Nacionales de Estadística y se pueden descargar gratuitamente tras completar unos breves formularios. Ocasionalmente, las encuestas están disponibles en los sitios web de los ministerios de salud y/o en los observatorios de uso de drogas, ya que son los responsables de recopilar estos datos. En caso de que no estén disponibles en línea, es aconsejable, en primer lugar, identificar la institución que realiza y recopila la encuesta y, en segundo lugar, ponerse en contacto directamente con la institución responsable.

Hay otras encuestas, que no se centran directamente en la salud o el consumo de drogas, que incluyen información sobre el consumo de tabaco por parte de los consumidores. La encuesta de ingresos y gastos es un ejemplo de ello. Esta encuesta, realizada cada cinco a diez años, se utiliza para estimar patrones de gasto representativos en varios países. El objetivo es estimar una canasta básica de bienes utilizada para calcular el índice de precios al consumidor. Recopilan una gran cantidad de características socioeconómicas de los hogares. Lo que es más importante, la encuesta recopila información sobre la cantidad y el valor de los bienes y servicios que los consumidores compran o reciben durante un período de tiempo. Esos períodos suelen variar según el tipo de bien/servicio (los bienes/servicios de consumo menos frecuente, como los gastos en educación, la compra de bienes duraderos, etc., tienen un período recordatorio más largo). Los productos de consumo frecuente, como los cigarrillos, tienen un período recordatorio relativamente corto (por ejemplo, la semana o el mes anterior, etc.). El consumo de cigarrillos en los hogares se utiliza para estimar el consumo de la población en general. A veces estas encuestas no recopilan información sobre las cantidades que se compran sino sobre lo que se gasta en los artículos. Para transformar esa información a cantidades que se compran, hay que considerar un precio promedio de los cigarrillos. Las advertencias para esta estrategia son similares a las de la transformación de los ingresos por impuestos ad valorem en las ventas de cigarrillos que pagan impuestos, que se han mencionado anteriormente.

Las encuestas de ingresos y gastos recopilan datos sobre los gastos de los hogares y los gastos individuales. En general, los datos están disponibles como dos conjuntos de datos separados. Para estimar el consumo de cigarrillos de la población, los investigadores deben utilizar el conjunto de datos de gastos individuales, que por lo general contiene las cantidades que se compran. El conjunto de datos sobre los gastos de los hogares suele contener información sobre los gastos en productos adquiridos por una persona (es decir, el jefe del hogar), pero no necesariamente consumidos por él (por ejemplo, alimentos, ropa, etc.), o sobre “bienes públicos” dentro de los hogares, como la electricidad, la calefacción, los gastos de vivienda, etc.

Una limitación de las encuestas de ingresos y gastos es que es probable que los menores de edad tergiversen su consumo de cigarrillos. Es probable que no informen o informen un consumo de cigarrillos subestimado con la intención de ocultarles su hábito a los padres.

Cuando los países carecen de los recursos humanos o el financiamiento necesarios para realizar ellos mismos encuestas nacionales sobre el consumo de cigarrillos, los organismos internacionales pueden recopilar esos datos en su lugar. La Encuesta Demográfica y de Salud (EDS) es financiada por la Agencia de los Estados Unidos para el Desarrollo Internacional (ahora realizada por ICF, *Inner City Fund*). Esta encuesta ha recopilado más de 400 encuestas en 93 países de ingreso mediano y bajo desde principios de la década de 1990. La encuesta se realiza normalmente cada tres o cinco años. Aunque no todos los países son encuestados por igual, la mayoría dispone de al menos dos encuestas. La encuesta EDS se centra en la salud materna e infantil. Recientemente, la EDS comenzó a recopilar otros aspectos relacionados con el tabaquismo en el hogar. La encuesta incorpora información sobre la frecuencia del consumo de cigarrillos y otros productos de tabaco entre hombres y mujeres de 15 a 49 años. Sin embargo, la información sobre la intensidad del tabaquismo solo está disponible para los hombres. Si bien se trata de una limitación, es una menor, ya que la prevalencia del tabaquismo es baja entre las mujeres en los países de ingreso mediano y bajo. Si se puede suponer razonablemente que la prevalencia (y la intensidad) del tabaquismo es constante a lo largo del tiempo, entonces la EDS es una fuente de datos viable.

La Encuesta Mundial de Tabaquismo en Adultos (GATS, por sus siglas en inglés) es otra fuente de datos sobre el consumo de tabaco en países de ingreso mediano y bajo. La encuesta está financiada por los Centros para el Control y la Prevención de Enfermedades de los Estados Unidos (CDC) como parte de su sistema global de vigilancia del tabaco (GTSS, por sus siglas en inglés). La GATS es una encuesta representativa de los hogares a

nivel nacional para la población de 15 años o más. El instrumento tiene preguntas sobre la frecuencia e intensidad del consumo de productos de tabaco. Los datos de la GATS son comparables a través del tiempo. Hasta ahora, los datos de la GATS están disponibles para 25 países. Sin embargo, no todos los países disponen de dos o más encuestas. Esto hace que el análisis de brechas no sea viable para los países que disponen de menos de dos encuestas.

4.4.4 Datos de población

Los organismos gubernamentales de estadística publican datos de población nacional de forma periódica. Comúnmente, estos datos se informan anualmente. Estos datos se incluyen en los datos del censo y son estimaciones oficiales del país. En la mayoría de los casos, las estimaciones son creíbles. Sin embargo, se aconseja precaución en el caso de los países que se sabe que se inmiscuyen en sus datos por razones políticas. Si hay dudas sobre la integridad de los datos, los investigadores pueden basarse en estimaciones independientes publicadas por organizaciones internacionales. Diversos organismos de las Naciones Unidas, entre ellos el Banco Mundial o la División de Población del Departamento de Asuntos Económicos y Sociales de las Naciones Unidas, disponen de estimaciones de población confiables y públicas. Las estimaciones de población alternativas se pueden utilizar como una comprobación de la solidez.

Si en el análisis se utilizan datos a nivel de población, estos se deben desglosar para garantizar que las estructuras de edad de la población y la prevalencia de tabaquismo sean similares. Los niños pequeños, que suelen estar representados en las estimaciones a nivel de población, no fuman y deben excluirse de las estimaciones a nivel de población. De lo contrario, la base de comparación entre la población y las estimaciones de prevalencia no es comparable.

Por ejemplo, si la prevalencia de tabaquismo en adultos se aplica a la población de 16 años o más, es necesario desglosar los datos de la población para tener en cuenta el mismo rango de edad. Si la prevalencia de tabaquismo entre los jóvenes es aplicable a la población de 12 a 16 años, entonces los datos de población deben reflejar también el mismo rango de edad.

En Sudáfrica, por ejemplo, la estimación de población para 2018 fue de 57.725.606 habitantes. Los jóvenes de 1 a 14 años alcanzaron los 17.043.517, lo que sugiere que la población adulta representaba 40.682.089 (StatsSA, 2018). Sería incorrecto aplicar una tasa de prevalencia del tabaquismo en adultos del 20% a la población total. Sin desagregación, esto sugeriría que hay 11.545.121 fumadores. Esto exageraría demasiado el número de fumadores adultos reales (8.136.418 fumadores). Si la intensidad del consumo de tabaco es de 10 cigarrillos por fumador al día, utilizando ambas estimaciones, el tamaño total del mercado se estimaría en 42.100 millones de cigarrillos y 29.700 millones de cigarrillos, respectivamente. Los investigadores sobreestimarían el tamaño verdadero del mercado en 12.400 millones de cigarrillos, una diferencia significativa. Por consiguiente, es necesario tener en cuenta la población representada en los datos cuando se utilicen dos conjuntos de datos diferentes.

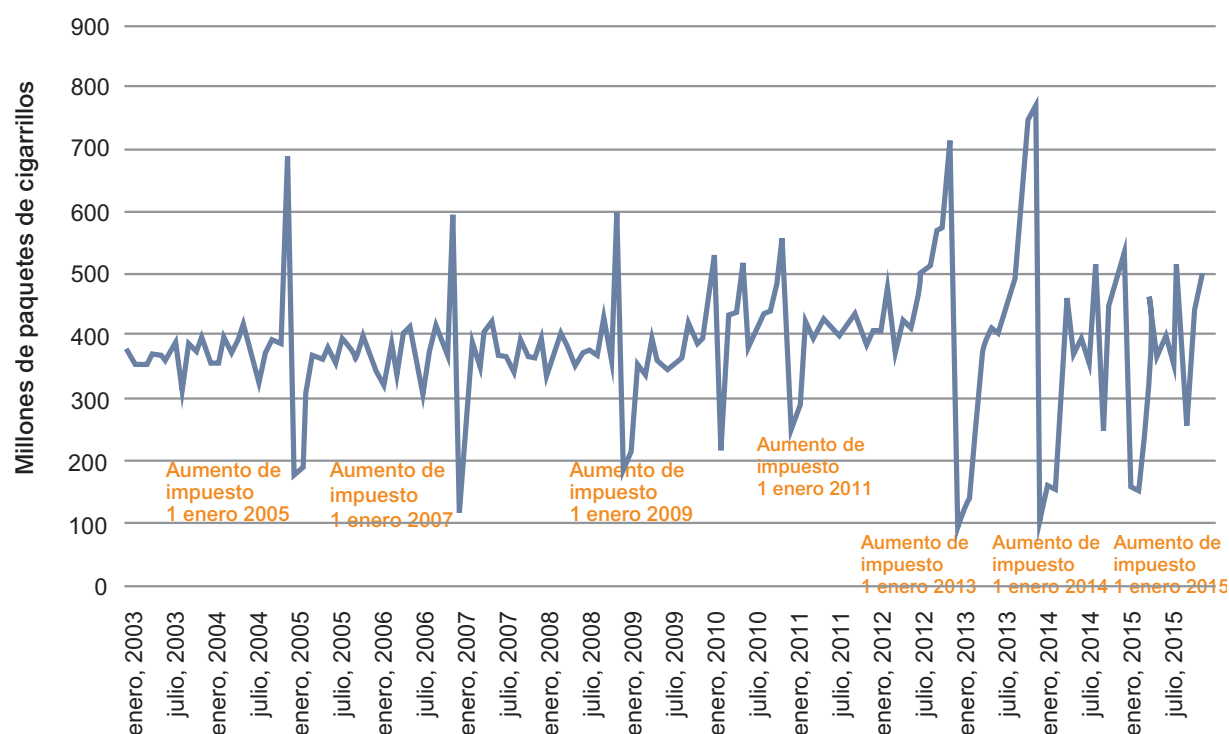
4.5 Desafíos con los datos

4.5.1 Anticipando cambios fiscales

La industria tabacalera anticipa estratégicamente (almacenando o surtiendo por adelantado) los cigarrillos cuando se espera o se anuncia un aumento de impuestos. La industria tabacalera inunda el mercado con cigarrillos antes de que el cambio de impuestos entre en vigor para evitar el pago del nuevo impuesto. Cabe esperar un aumento temporal de las ventas, seguido de una caída repentina de las ventas de tabaco después de la aplicación de los impuestos. El Gráfico 11 muestra este fenómeno en las Filipinas.

La anticipación puede comprometer el uso del método de brecha, ya que el aumento repentino (disminución) en las ventas que pagan impuestos puede atribuirse erróneamente a una disminución (aumento) en el comercio ilícito, cuando la fluctuación en las ventas es solo un proceso artificial propiciado por la industria. Si se produce

Gráfico 11 Volumen de retiro de cigarrillos en Filipinas (producción nacional)



Origen: Ross *et ál.*, (2017)

una anticipación, el método de brecha se debe usar con cuidado. Las tendencias de venta a largo plazo se deben utilizar para suavizar los patrones de consumo erráticos. Los promedios de series cronológicas, lo suficientemente largos como para compensar la anticipación, se deben usar como estimaciones totales del mercado. Alternativamente, los períodos con un crecimiento/disminución no natural en las ventas (es decir, más del 10%, 20%, etc.) se pueden excluir de los cálculos. Para el método de brecha, la clave es la evolución del mercado con impuestos pagados, más que las cantidades reales vendidas.

4.5.2 Inconsistencia en los períodos de tiempo

El método de brecha requiere que el período de ventas con impuestos pagados y el período de consumo sean iguales. Sin embargo, las estimaciones del consumo se suelen recopilar durante un período de tiempo limitado (por ejemplo, un mes específico). Las comparaciones realizadas entre los datos de consumo y de ventas mensuales pueden pasar por alto el efecto de las fluctuaciones de las ventas estacionales de los cigarrillos que pagan impuestos. Lo ideal sería que los datos de consumo se recopilaran para un año, lo que rara vez ocurre. Los investigadores deben asumir que los patrones de consumo son constantes durante el año, aunque solo se disponga de datos mensuales. Esta puede ser una suposición sólida, ya que los efectos de las ventas estacionales sugerirían que el consumo es mayor en algunos meses que en otros. Pero sin esta suposición necesaria, no se puede utilizar el método de brecha. Los datos de las ventas deberían tener al menos un año.

4.5.3 Subregistro de la prevalencia e intensidad de tabaquismo

El subregistro del consumo de cigarrillos es un problema. Sin embargo, no es inusual que las estimaciones del consumo total sean inferiores a las ventas con impuestos pagados. Esto se debe a que algunos grupos de población (poblaciones institucionalizadas, por ejemplo) pueden no estar incluidos en la muestra, o a que simplemente no se dispone de esa información (ventas libres de impuestos). Sin introducir sesgos en la muestra, se puede suponer que este tipo de omisiones son constantes y pueden ignorarse razonablemente. El mayor problema surge cuando las personas de la muestra subestiman la información sobre la prevalencia y la intensidad de tabaquismo, porque este tipo de omisión no es aleatoria. El subregistro de la prevalencia se produce cuando los fumadores declaran erróneamente si fuman. El subregistro de la intensidad se produce cuando declaran erróneamente la cantidad consumida en el período relevante (por ejemplo, informan una cantidad menor que la que realmente fumaron).

Se desconoce la magnitud del subregistro. Algunos autores han documentado su frecuente existencia comparando diferentes fuentes de datos de usuarios (es decir, dos encuestas de consumidores diferentes) (Liber y Warner, 2018; Graham y Owen, 2003) o cuando comparan las respuestas de los usuarios con el nivel de cotinina u otros químicos (Stelmach *et ál.*, 2015). Un metaanálisis determinó que la exactitud de la autoevaluación depende del entorno de la encuesta, la población de estudio, los métodos de medición y el propósito del estudio (Patrick *et ál.*, 1994).

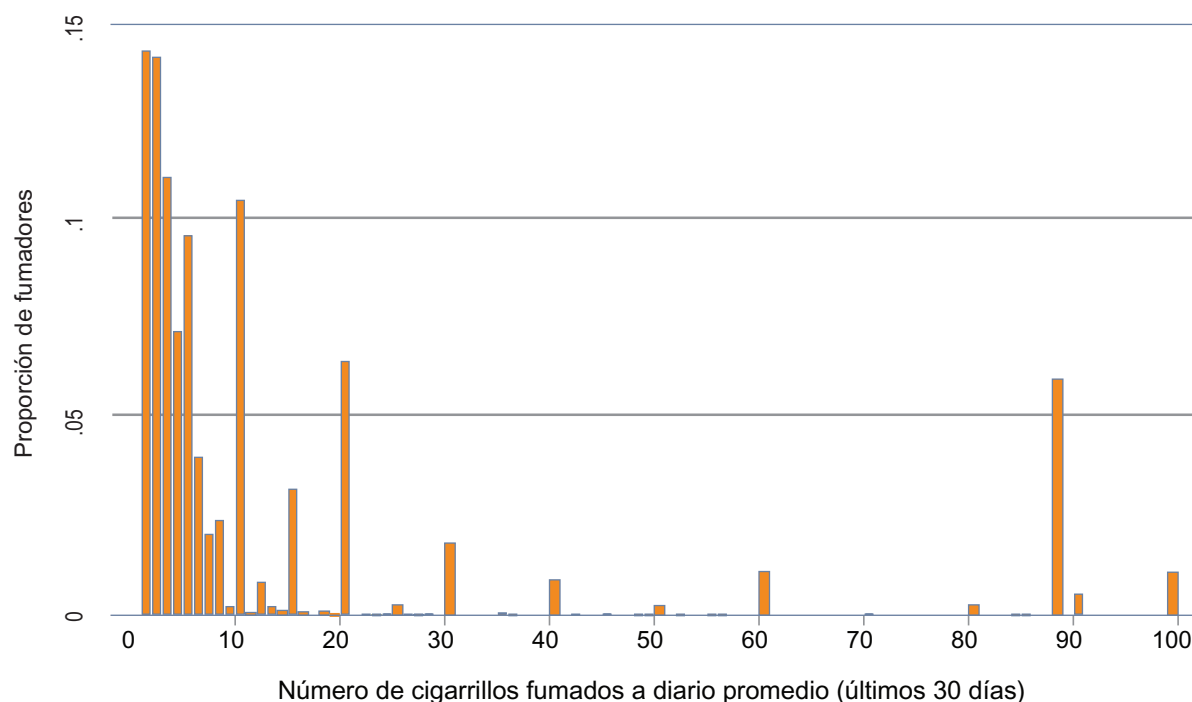
El método de brecha se puede seguir utilizando cuando se presenta un subregistro. El aspecto más importante es qué tan realista es la suposición del subregistro constante a lo largo del tiempo. Puede resultar difícil comprobar esta suposición, ya que los datos para la comparación son escasos o inexistentes. En el caso de los Estados Unidos, se ha demostrado que en un contexto de disminución de la prevalencia, el subregistro de la prevalencia y la intensidad se mantuvo bastante constante a lo largo del tiempo (Liber y Warner, 2018).

4.5.4 Sesgo en el periodo recordatorio de la intensidad de tabaquismo

La intensidad de tabaquismo se mide preguntándoles a los encuestados cuántos cigarrillos fumaron en un período de tiempo anterior, generalmente un día o una semana, mientras que en otras encuestas se les pide a los encuestados que calculen el número promedio de cigarrillos que se fuman cada día o semana. Por ejemplo, el modelo de Preguntas para Encuestas sobre el Tabaco de la Organización Mundial de la Salud recomienda la siguiente pregunta: “En promedio, ¿cuántos de los siguientes productos fuma actualmente cada (día/semana)?”.

Es probable que los encuestados proporcionen una estimación de su consumo y a menudo están sujetos a subestimar la información (véase subsección anterior). Los investigadores no pueden validar la respuesta autoevaluada. Investigaciones anteriores sugieren que el número de cigarrillos consumidos tiende a agruparse en decenas (por ejemplo, 10, 20, 30, etc., cigarrillos por día). Esto se conoce como sesgo en el periodo recordatorio. Una explicación probable es que los consumidores de cigarrillos pueden pensar más en términos de tamaños de cajetillas (20 cigarrillos en una cajetilla) que en el número de cigarrillos individuales que consumieron. Los fumadores también pueden redondear su respuesta a las decenas por conveniencia. Estos sesgos en el periodo recordatorio introducen inexactitudes en los datos de intensidad de tabaquismo. El Gráfico 12 muestra las respuestas de distribución de la intensidad de tabaquismo para la Encuesta de Población General de Chile de 2016 sobre el consumo de drogas (Observatorio Chileno de Drogas, 2017). Los picos son claramente visibles en 10, 20, 30 y 40 cigarrillos (los tamaños de cajetillas en Chile vienen en 10 y 20 cigarrillos). El sesgo en el periodo recordatorio no sugiere que los encuestados estén tergiversando intencionadamente la verdad, sino simplemente que los consumidores tienden a agrupar sus respuestas de la manera más fácil de recordar. Lamentablemente, no existe una forma satisfactoria de abordar este problema, aunque debería reconocerse debidamente como una limitación de los datos.

Gráfico 12 Histograma de la intensidad de tabaquismo en la población general de Chile, 2016



Origen: Estimaciones de los autores del Observatorio Chileno de Drogas (2017)

4.6 Estimación de la brecha y presentación de los datos

El primer paso es obtener el consumo total informado a partir de encuestas. Luego se puede utilizar el método de brecha para estimar la diferencia entre las ventas de cigarrillos con impuestos pagados y el consumo total. Cuando se utilizan encuestas a nivel de población, es necesario identificar primero la estructura de la muestra. Esto es importante porque es necesario considerar la representación relativa de cada grupo de edad. Por lo general, los estudios utilizan diseños de encuestas en varias etapas en el proceso de recopilación de datos. La documentación proporcionada junto con los datos de la encuesta suele ofrecer información detallada sobre el diseño de muestreo específico que se utilizó. Los paquetes estadísticos, como Stata, pueden manejar tales diseños de encuestas para hacer inferencias precisas sobre la población.

El comando de Stata `<svyset>` declara que se está utilizando un diseño de encuesta. El comando permite definir el número de etapas de muestreo, métodos y ponderaciones que se utilizaron en el estudio. Los estratos, las unidades primarias de muestreo (UPM) o los conglomerados se pueden definir fácilmente. Los estratos son unidades geográficas amplias en las que se ubican las UPM y definen la representatividad geográfica de las encuestas. Las UPM son unidades geográficas en las que se recopilan datos y que se definen y escogen, generalmente de forma aleatoria. Las ponderaciones de muestreo dan cuenta de la importancia original de cada subgrupo de la población dentro de la población total, dada la estructura de muestreo. En otras palabras, cada observación de la muestra representa a muchos individuos/hogares de la población, y esa relación se resume mediante las ponderaciones de la muestra.

Por ejemplo, el Décimo Segundo Estudio Nacional de Drogas en Población General de Chile utiliza un diseño de muestreo en tres etapas. En primer lugar, considera las UPM basadas en los datos del censo. Tales UPM son un grupo de bloques cercanos (manzanas) dentro de 108 comunas urbanas en Chile. La selección de la UPM es

probabilística. A las UPM se les asignan probabilidades basadas en el tamaño relativo de cada comuna. La segunda etapa es la selección de los hogares dentro de las UPM. El número de hogares dentro de cada UPM se asigna en función de la proporción de hogares dentro de la UPM. En la tercera etapa, los encuestados (de 12 a 65 años) se seleccionan al azar dentro de cada hogar.

El comando `<svyset>` tiene una secuencia de pasos. Es aconsejable borrar cualquier diseño anterior escribiendo `<svyset, clear>`. Segundo, la sintaxis de `<svyset>` es: `<svyset psu [weight] [, design options] [| ssu, design options] ... [options]>` donde `psu` es el nombre de una variable que identifica la UPM en los datos, `weight` identifica la ponderación de muestreo, `ssu` identifica las unidades de muestreo en la segunda etapa, y así sucesivamente. Tercero, una vez que se declara el diseño de la muestra, el prefijo `<svy:>` se requiere con cada comando subsiguiente para dar cuenta del diseño de la muestra.

Supongamos que se realizó un estudio de datos con un diseño de dos etapas. La variable designada para indicar el conglomerado de hogares en una zona geográfica pequeña es `psu`. La variable de ponderación, `weight`, da la importancia relativa dentro de la población de las personas seleccionadas en la segunda etapa. Por lo tanto, el comando para establecer el diseño de la muestra es:

```
svyset psu [pw=weight]
```

Supongamos también que en esa encuesta la variable `q10` recoge las respuestas a la pregunta “¿Ha fumado en los últimos 30 días?”, con 0 = no; 1 = sí; 2 = no sabe/no responde. Para construir una variable que señale si la persona es un fumador habitual (habiendo fumado en el mes anterior) el código de Stata sería:

```
gen smoker=(q10==1)  
replace smoker=. if smoker > 1
```

Las personas que no quisieron identificar su condición de fumadores no serán registradas ni como fumadores ni como no fumadores. Para obtener la prevalencia mensual de tabaquismo en la población, se puede utilizar el siguiente comando:

```
svy: mean smoker
```

Esto da la proporción de la población que fuma regularmente. Para obtener el número real de fumadores en la población, se puede utilizar el siguiente comando:

```
summarize weight
```

```
. su weight
```

<i>variable</i>	<i>obs</i>	<i>Mean</i>	<i>Std. Dev.</i>	<i>Min</i>	<i>Max</i>
<i>weight</i>	<i>41,392</i>	<i>554.0998</i>	<i>1144.936</i>	<i>2</i>	<i>32745</i>

La columna “Obs” da el número de observaciones en el conjunto de datos, y la columna “Mean” da el valor promedio de la ponderación en el conjunto de datos. Multiplicando ambas columnas se obtiene el número de personas en la muestra. Multiplicando este número por la proporción de fumadores se obtiene el número de fumadores habituales de la muestra.

Suponiendo que la variable q11 reúne las respuestas a la pregunta “¿Cuántos cigarrillos ha fumado diariamente en los últimos 30 días?” entonces

```
svy: mean q11 if smoker==1
```

representa la intensidad promedio diaria por fumador. Multiplicando esta intensidad promedio por 365 se obtiene la cantidad total de cigarrillos consumidos en un año por fumador. (Si bien casi todas las encuestas estiman la intensidad diaria, también se pueden presentar como semanales o mensuales. En estos casos, los investigadores multiplicarían por 52 o 12, respectivamente). A su vez, al multiplicar el total de cigarrillos fumados por fumador-año por el número de fumadores se obtiene el número total de cigarrillos consumidos en un año.

Alternativamente, el código de abajo genera una nueva variable:

```
summarize q11 if smoker==1 [w=weight]  
gen total = r(mean)*365
```

La primera línea proporciona la intensidad promedio diaria del consumo de tabaco, mientras que la segunda línea crea una nueva variable con el número total de cigarrillos fumados por persona-año.

En algunos casos, hay información sobre fumadores poco frecuentes que no se reúne mediante preguntas como las presentadas anteriormente. Es raro que las personas que fuman 10 cigarrillos en un mes reporten “0,33” cuando se les pregunta sobre su consumo diario promedio. Para ellos, puede haber información sobre el número de días que han fumado (“En el último mes, ¿cuántos días ha fumado?”) y sobre la cantidad fumada por día (“En los días que ha fumado durante el último mes, ¿cuántos cigarrillos ha fumado en promedio?”). Así que, para una persona que informa haber fumado 5 días en el último mes y haber fumado 3 cigarrillos en promedio en esos días, el número de cigarrillos que fumó en un año sería 180 (3 * 5 * 12). Si se dispone de esa información, esos consumos se pueden sumar fácilmente para obtener el número total de cigarrillos fumados por esas personas en un año.

Suponiendo que la variable q12 reúne información sobre el número de días que fumó en el último mes, mientras que la q13 recoge información sobre el número de cigarrillos fumados en promedio en esos días, entonces los códigos para estimar el consumo anual total son:

```
summarize q12 if smoker==1 [w=weight]  
gen days = r(mean)  
summarize q13 if smoker==1 [w=weight]  
gen cigs = r(mean)  
gen total = days*cigs*12
```

Los datos de las encuestas suelen estar sujetos a errores estadísticos. Es aconsejable establecer intervalos de confianza en torno a las estimaciones del consumo total mediante el bootstrapping de la muestra:

```
bootstrap sumboot=r(sum), reps(1000) force cformat(%9.0f): sum total [fw=weight]
```

El *bootstrap* es un método no paramétrico que consiste en tomar al azar (con reemplazo) una serie de submuestras de la población del estudio (en la línea anterior, esa cifra es de 1.000). Se construyen mil muestras artificiales utilizando los datos originales (mil son lo suficientemente grandes como para producir errores estándar confiables). Se pueden calcular las estadísticas solicitadas (número de cigarrillos fumados en una población en un año) junto con intervalos de confianza del 95% para las estadísticas.

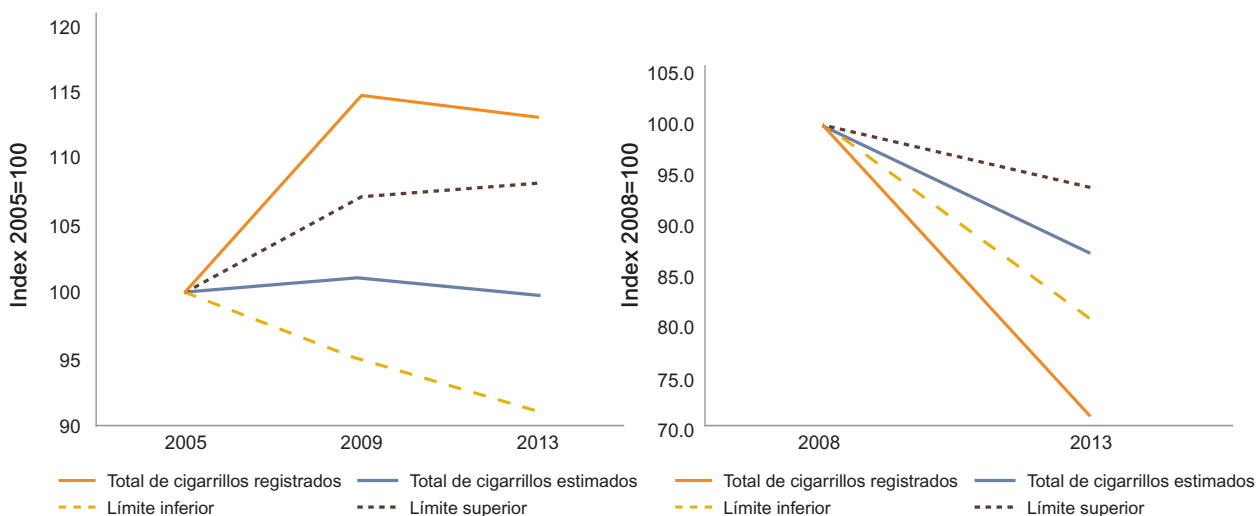
Estos cálculos se deben hacer para cada año de la encuesta. Una vez completado, el análisis de brechas se puede estimar utilizando solo tres puntos de datos (el mercado total y sus límites de intervalo de confianza del 95%) y las ventas de cigarrillos con impuestos pagados. Debido al subregistro, es probable que el consumo informado o las ventas totales del mercado sean más bajas que las ventas con impuestos pagados, lo que sería un resultado sin sentido (lo que sugiere que el comercio ilícito es negativo).

El método de brecha estima la evolución del mercado que no paga impuestos a lo largo del tiempo, en el que se prevé que el mercado ilícito sea su principal componente. Por lo tanto, debe definirse un año base, y la evolución de los diferentes mercados se puede estimar a partir del año base. El mercado total (y sus intervalos de confianza) y el mercado de impuestos pagados se pueden indexar a 100 para el año base. Los valores del otro año se pueden imputar usando la simple regla de tres. Por lo tanto, si el mercado con impuestos pagados para el año base es de 1.500 millones de cigarrillos y el consumo estimado para el mismo año es de 1.300 millones, se puede suponer que ambas cifras son iguales a 100. Si al año siguiente, el mercado con impuestos pagados y el consumo estimado son 1.600 millones y 1.350 millones, respectivamente, entonces, en términos del año base, equivalen a 106,7 (1.600 millones * 100 / 1.500 millones) y 103,8 (1.350 millones * 100 / 1.300 millones), respectivamente.

El Gráfico 13 muestra el análisis de la brecha entre Brasil y Argentina. Para Argentina (izquierda), la evolución de las ventas con impuestos pagados se encuentra por encima de los límites del intervalo de confianza del 99% del mercado total. Esto sugiere que la tasa de crecimiento del mercado lícito fue mayor que la tasa de crecimiento total del mercado. El mercado ilícito disminuyó en términos proporcionales bajo la suposición del subregistro invariante en el tiempo. Es probable que el mercado ilícito se haya reducido, ya que la prevalencia del tabaquismo disminuyó durante ese período.

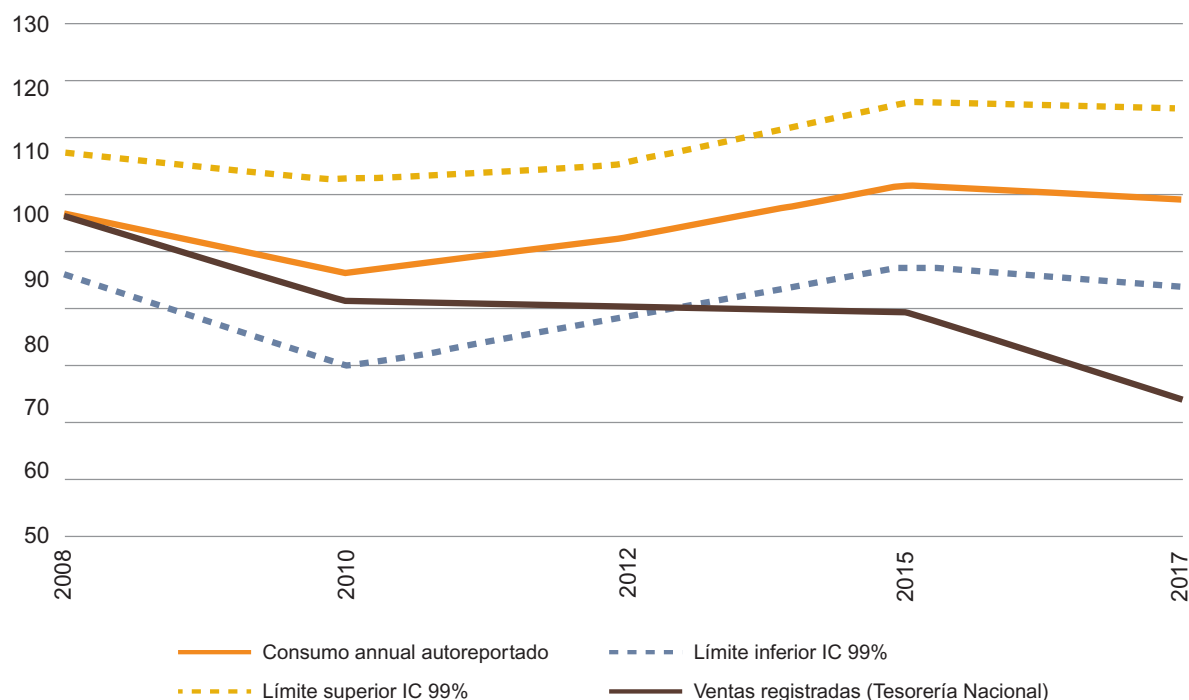
En el caso de Brasil (derecha), la tasa de crecimiento del mercado con impuestos pagados disminuyó en relación con la tasa de crecimiento del mercado total. La tasa de crecimiento del mercado con impuestos pagados está por debajo del intervalo de confianza del 99%. Este resultado implica que el mercado que no paga impuestos creció en términos relativos. La cifra sugiere que el total de cigarrillos consumidos disminuyó en el período. Sin embargo, las ventas legales también disminuyeron. El cambio relativo en el tamaño del mercado ilícito normalmente no se podría determinar en tal caso. En el caso concreto de Brasil, hay fuertes indicios de que el tamaño del mercado ilícito creció, ya que el consumo indicado por la encuesta en 2013 fue mayor que las ventas con impuestos pagados.

Gráfico 13 Análisis de la brecha para Argentina (izquierda) y Brasil (derecha)



Origen: Paraje (2018):

Gráfico 14 Análisis de brechas para Sudáfrica (2002 = 100)



Origen: Vellios *et ál.*, (2019).

En el Gráfico 14 se muestra el análisis de brechas para Sudáfrica. Al igual que en el caso de Brasil, el mercado registrado se situó por debajo del intervalo de confianza del 99%, lo que indica un aumento del mercado ilícito de cigarrillos. A diferencia del caso de Brasil, en Sudáfrica ese aumento se produjo con una prevalencia estadísticamente constante del tabaquismo.

4.7 Casos prácticos

4.7.1 El caso de Canadá

Un estudio en Canadá investigó las tendencias del contrabando de cigarrillos mediante un análisis de brechas, centrándose especialmente en las provincias de Ontario y Quebec (Guindon *et ál.*, 2016). El análisis de brechas contrasta las estimaciones de cigarrillos con impuestos pagados con los datos de consumo de las encuestas sobre el comportamiento de los fumadores. Los autores del estudio utilizan las cifras de cigarrillos con impuestos pagados que se han reportado al Ministerio de Salud de Canadá. Además, los autores tienen en cuenta dos grandes encuestas nacionales, la Encuesta de Salud de la Comunidad Canadiense (CCHS, por sus siglas en inglés) y la Encuesta sobre el Tabaco, el Alcohol y las Drogas de Canadá (CTADS, por sus siglas en inglés), para estimar el consumo de cigarrillos. Los autores determinaron la relación entre el consumo de cigarrillos reportado por la CCHS y la CTADS y las ventas de cigarrillos con impuestos pagados. Suponiendo que el patrón de subregistro es invariante en el tiempo, si la proporción es constante (aumenta) implica una proporción constante (aumento) de cigarrillos de contrabando. Estas proporciones se calcularon de 1999 a 2013 con una frecuencia semestral.

Los análisis mostraron “relativamente pocos cambios a principios de la década de 2000, interrumpidos por un claro salto al alza alrededor de 2005, seguido de una tendencia decreciente desde aproximadamente 2007 a 2009” (Guindon *et ál.*, 2016). Los autores también encontraron que las estimaciones para Ontario y Quebec siguieron tendencias similares, pero con mayores magnitudes. Los investigadores pudieron determinar qué partes de las tendencias eran significativas debido a la inclusión de intervalos de confianza.

4.7.2 Análisis de brechas para cinco países de América Latina

En un estudio reciente se evaluó la evolución de la cuota de cigarrillos ilícitos en Argentina, Brasil, Chile, Colombia y Perú (Paraje, 2018). Las ventas de cigarrillos con impuestos pagados se compararon con las estimaciones de consumo de las encuestas representativas a nivel nacional. Los datos sobre las ventas de cigarrillos con impuestos pagados se obtuvieron de fuentes oficiales (Argentina, Brasil y Chile) o de estimaciones de empresas consultoras (Colombia y Perú). En Argentina se utilizaron dos encuestas: La Encuesta Nacional de Factores de Riesgo (ENFR) y la Encuesta Nacional sobre la Prevalencia del Consumo de Sustancias Psicoactivas (ENPRECOSP). Las encuestas tienen diferentes marcos de muestreo. La ENFR recopila datos sobre la población de 18 años y más, mientras que la ENPRECOSP recopila datos sobre la población de 16 a 65 años. En este estudio se estimaron dos tendencias diferentes del comercio ilícito sobre la base de cada conjunto de datos.

Los investigadores encontraron que el comercio ilícito en Argentina disminuyó entre 2005 y 2009 y se mantuvo constante hasta 2013. En el caso de Brasil, hubo cambios estadísticamente significativos en la diferencia entre las estimaciones de las tendencias de consumo y las ventas con impuestos pagados entre 2008 y 2013. Esto implicó un aumento de la cuota del contrabando de cigarrillos. Otros estudios brasileños, que utilizaron la misma metodología, confirmaron estos resultados (Szklo *et ál.*, 2017; Iglesias *et ál.*, 2016).

El estudio chileno no encontró ninguna diferencia estadísticamente significativa entre los cigarrillos con impuestos pagados y las tendencias de consumo (población de 12 a 65 años) para el período 2008 a 2014. Esto implica que la cuota del comercio ilícito chileno no cambió durante este período.

El estudio de Colombia no encontró ninguna diferencia estadísticamente significativa entre los cigarrillos con impuestos pagados y las tendencias de consumo (población de 12 a 65 años) para 2008 a 2013. Esto implica que la cuota del comercio ilícito colombiano no cambió durante este período.

Por último, el estudio peruano no encontró diferencias estadísticamente significativas entre los cigarrillos con impuestos pagados y las tendencias de consumo (población de 12 a 65 años) para 2006 a 2010. Esto implica que la cuota del comercio ilícito peruano no cambió durante este período.

4.7.3 Estimación del comercio ilícito en Sudáfrica

En un estudio realizado en Sudáfrica se estimaron las ventas de cigarrillos con impuestos pagados utilizando los ingresos fiscales y el consumo. El estudio tenía por objeto estimar la prevalencia y el tamaño y la evolución del comercio ilícito (Blecher, 2010). En el estudio se pudo estimar el número de cigarrillos con impuestos pagados porque Sudáfrica grava los cigarrillos con un impuesto especial al consumo. Dividiendo los ingresos anuales por concepto de impuestos al tabaco por la cantidad de impuestos específicos unitarios se obtiene el número total de cigarrillos lícitos. Los datos sobre la prevalencia del tabaquismo se obtuvieron de la Encuesta sobre toda la Media y los Productos (AMPS, por sus siglas en inglés). Los datos sobre la intensidad de tabaquismo provenían de tres fuentes de datos: investigaciones independientes anteriores, datos de la AMPS y estimaciones de empresas consultoras privadas. La población total de fumadores se simuló utilizando las estimaciones de población total de la Oficina de Estadísticas de Sudáfrica. Mediante la simulación de diferentes niveles de subregistro del consumo (del 5% al 10% del consumo reportado), el autor proporcionó no solo las tendencias, sino también las cifras de la penetración del mercado de cigarrillos ilícitos.

El autor constató que la penetración del mercado del comercio ilícito creció rápidamente entre 1997 y 2000, situándose entre el 9,4% y el 11,5% del mercado total. El mercado disminuyó lentamente entre 2000 y 2007.

En un estudio reciente, Vellios *et ál.* (2019) produjeron la tendencia y el tamaño del mercado ilícito de cigarrillos en el período comprendido entre 2002 y 2017. Utilizando una metodología similar a la de Blecher (2010), Vellios *et ál.* (2019) estimaron el consumo de cigarrillos a partir de la AMPS (2002-2014) y el Estudio de la Dinámica del Ingreso Nacional (NIDS, por sus siglas en inglés) para 2008, 2010, 2012, 2015 y 2017. Los autores asumieron que el subregistro de fumadores en estas encuestas es del 5% y 10% para la AMPS, y del 15% y 20% para el NIDS. El NIDS mostró consistentemente una menor prevalencia de tabaquismo que la AMPS, razón por la cual se utilizó un porcentaje mayor de subregistro en estos cálculos. Los investigadores descubrieron que el mercado ilícito de cigarrillos aumentó bruscamente desde 2009, alcanzando entre el 30% y el 35% del mercado total en 2017.

4.7.4 Medición de las brechas fiscales en el Reino Unido

El servicio de HM Revenue & Customs (Rentas y Aduanas de Su Majestad, HMRC, por sus siglas en inglés), la autoridad fiscal del Reino Unido proporciona regularmente estimaciones sobre la evasión de impuestos en el Reino Unido. Estas estimaciones se proporcionan para el tabaco y varios otros bienes, como el alcohol y los aceites de hidrocarburos, utilizando análisis de brechas (HM Revenue & Customs, 2019). HMRC incluye los cigarrillos y el tabaco para armar a mano en sus cifras para el tabaco.

En la edición de 2019, las estimaciones de cigarrillos se calcularon utilizando las estimaciones de prevalencia y consumo de tabaco proporcionadas por la Encuesta General de Estilo de Vida (GLF, por sus siglas en inglés), la Encuesta de Opiniones y Estilo de Vida (OPN, por sus siglas en inglés) y la Encuesta de Salud para Inglaterra (HSE, por sus siglas en inglés). Las estimaciones de la población adulta se extrajeron de la Oficina de Estadísticas Nacionales (ONS, por sus siglas en inglés). También se utilizó un subregistro de consumo de cigarrillos o “factor de aumento”. Este factor “se calculó tomando las estimaciones del consumo total de la GLF en un año base, y comparándolo con el consumo basado en las declaraciones reales al HMRC y una estimación de los cigarrillos comprados legítimamente en el extranjero”. Los intervalos de confianza se estimaron para tener en cuenta las imprecisiones del proceso de encuesta. En este caso particular, los intervalos de confianza son especialmente importantes porque los fumadores duales (aquellos que fuman tanto cigarrillos manufacturados como cigarrillos armados a mano) se deben contabilizar adecuadamente. La metodología para el tabaco para armar sigue una metodología similar a la de los cigarrillos manufacturados.

El HMRC estimó que el consumo de cigarrillos con impuestos pagados ha “disminuido constantemente de 49.500 millones de cigarrillos en 2005–06 a 27.500 millones de cigarrillos en 2017–18”, y “(desde) 2010–11 la estimación central del mercado ilícito ha sido bastante estable, oscilando entre 3.000 y 5.500 millones de cigarrillos” (HM Revenue & Customs, 2019). Esto implica una cuota del mercado ilícito del 9% del mercado total en 2017–18.

En el caso del tabaco para armar, la situación es diferente, ya que “el volumen de consumo de tabaco para armar con impuestos pagados aumentó de forma constante entre 2005–06 y 2013–14 y se ha mantenido relativamente estable desde entonces”, mientras que “el volumen del mercado ilícito de tabaco para armar ha mostrado una disminución a largo plazo desde 2005–06 y se estimó que será de 3,1 millones de kilogramos (toneladas) en 2017–18”. En este caso, el mercado ilícito representa el 32% del mercado total.

4.8 Conclusión

El análisis de brechas es un método fácil de aplicar, rápido y económico que proporciona estimaciones de las tendencias subyacentes del comercio ilícito de cigarrillos. Aunque no puede proporcionar estimaciones del nivel del mercado ilícito en un momento dado, les permite a los investigadores inferir si el comercio ilícito ha aumentado, se ha mantenido constante o ha disminuido en el tiempo.

En este capítulo se han presentado las principales necesidades de datos y los desafíos que rodean a este método. Este método no es particularmente intensivo en datos, pero se requieren por lo menos dos series de datos: las ventas de cigarrillos con impuestos pagados y el consumo durante al menos dos puntos de tiempo. El consumo se puede estimar mediante encuestas a consumidores, que suelen ser recopiladas por los organismos

oficiales. El otro elemento clave, la venta de cigarrillos con impuestos pagados, también suele estar disponible a bajo o nulo costo.

La preparación de los datos y la interpretación de los resultados pueden presentar algunos desafíos, y la mayoría de estos se abordan en el presente capítulo mediante ejemplos empíricos.

En general, el método ofrece un diagnóstico rápido de la tendencia del comercio ilícito de cigarrillos en un país y se puede utilizar para proporcionarles a las autoridades y al público en general una estimación rápida, sólida y fácil de entender sobre la tendencia. Como ocurre con todos los métodos, hay que considerar las limitaciones para evitar una interpretación excesiva (por ejemplo, dando estimaciones del tamaño del mercado).

5

Bibliografía

Adams, M., & Effertz, T. (2011). *Tabaksteuern, Schmuggel und unversteuerte Zigaretten – Zur Glaubwürdigkeit der “Entsorgungsstudie” der Tabakindustrie* [Impuestos especiales sobre el consumo de tabaco y el problema del contrabando – Relativo a la credibilidad del “Estudio sobre cajetillas de cigarrillos desechadas” de la industria tabacalera]. *Gesundheitswesen*, 73, 705–12.

Barker, D. C., Wang, S., Merriman, D., et ál. (2016). *Estimating cigarette tax avoidance and evasion: evidence from a national sample of littered packs*. *Tobacco Control*, 25, i38–43.

Blecher, E. (2010). *A mountain or a molehill: is the illicit trade in cigarettes undermining tobacco control policy in South Africa?* *Trends in Organized Crime*, 13(4), 299–315.

Bostoen, K., & Chalabi, Z. (2006). *Optimization of household survey sampling without sample frames*. *Int J Epidemiol*, 35, 751–755.

Centros para el Control y Prevención de Enfermedades. (Consultado en su versión en inglés el 11 de junio de 2019). *Global Tobacco Surveillance System Data*. <https://www.cdc.gov/tobacco/global/gtss/gtssdata/index.html>

Consroe, K., Kurti, M., Merriman, D., et ál. (2016). *Spring Breaks and Cigarette Tax Noncompliance: Evidence from a New York City College Sample*. *Nicotine Tob Res Off J Soc Res Nicotine Tob*, 18, 1773–9.

Gallagher, A. W. A., Evans-Reeves, K. A., Hatchard, J. L., et ál. (2019). *Tobacco industry data on illicit tobacco trade: a systematic review of existing assessments*. *Tobacco Control*, 28, 334.

Carvalho Figueiredo, V., Szklo, A. S., Iglesias R. M., et ál. (En curso). *Measuring illicit cigarette trade in Brazil*.

Drope, J., Schluger, N., Cahn, Z., et ál. (2018). *The Tobacco Atlas* (6.ª ed.). Atlanta: la Sociedad Americana del Cáncer y Vital Strategies 2018.

Comisión europea. (Consultado en su versión en inglés el 21 de agosto de 2019). *Special Eurobarometer 443: Public perception of illicit tobacco trade*. Bruselas, Bélgica: 2016.
<http://ec.europa.eu/commfrontoffice/publicopinion/index.cfm/ResultDoc/download/DocumentKy/73943>

Parlamento Europeo y el Consejo de la Unión Europea. (Consultado en su versión en inglés el 28 de agosto de 2014). *Directive 2014/40/EU of the European Parliament and the Council of 3 April 2014 on the approximation of the laws, regulations and administrative provisions of the Member States concerning the manufacture, presentation and sale of tobacco and related products and repealing Directive 2001/37/EC*.
http://ec.europa.eu/health/tobacco/docs/dir_201440_en.pdf

- Comisión europea. (Consultado en su versión en inglés el 6 de agosto de 2019). *Health Warnings*. https://ec.europa.eu/health/tobacco/law/pictorial_en
- Fundación Interamericana del Corazón – Argentina (FIC Argentina). (Manuscrito en curso). *Littered pack data to estimate illicit cigarette trade in in the city of Buenos Aires*.
- Frankel, M. (1983). *Handbook of Survey Research: Sampling Theory*. Nueva York: Academic Press, Inc.
- Gilmore, A. B., Rowell, A., Gallus, S., et ál. (2014). *Towards a greater understanding of the illicit tobacco trade in Europe: a review of the PMI-funded “Project Star” report*. *Tobacco Control*, 23, e51–61.
- Gilmore, A. B., Fooks, G., Drope, J., et ál. (2015). *Exposing and addressing tobacco industry conduct in low-income and middle-income countries*. *The Lancet*, 385, 1029–43.
- Graham, H., & Owen, L. (2003). *Are there socioeconomic differentials in under-reporting of smoking in pregnancy?* *Tobacco Control*, 12(4), 434.
- Guindon, G. E., Burkhalter, R., & Brown, K. S. (2016). *Levels and trends in cigarette contraband in Canada*. *Tobacco Control*.
- Guindon, G. E., Paraje, G. R., & Chávez, R. (2017). *Prices, inflation, and smoking onset: The case of Argentina*. *Economic Inquiry*.
- IARC. *IARC Handbooks of Cancer Prevention, Tobacco Control, Vol. 14: Effectiveness of Tax and Price Policies for Tobacco Control*. Lyon: Centro Internacional de Investigaciones sobre el Cáncer, 2011.
- Kurti, M., He, Y., Silver, D., et ál. (2019). *Presence of Counterfeit Marlboro Gold Packs in Licensed Retail Stores in New York City: Evidence from Test Purchases*. *Nicotine Tob Res*, 21, 1131–4.
- Kelley, K., Clark, B., Brown, V., et ál. (2003). *Good practice in the conduct and reporting of survey research*. *Int J Qual Health Care*, 15, 261–6.
- HM Revenue & Customs (2019). *Measuring tax gaps 2019 edition: Tax gap estimates for 2017–18*. Londres, Reino Unido: HM Revenue & Customs (Rentas y Aduanas de Su Majestad, HMRC, por sus siglas en inglés).
- Iglesias, R. M., Szklo, A. S., Souza, M. C., de Almeida, L. M. (2016). *Estimating the size of illicit tobacco consumption in Brazil: findings from the global adult tobacco survey*. *Tobacco Control*.
- International Tobacco Control (ITC) Policy Evaluation Project. (Consultado en su versión en inglés el 8 de septiembre de 2019). *About ITC*. 2017. <https://itcproject.org/about>
- John, R. M., & Ross, H. (2018). *Illicit cigarette sales in Indian cities: findings from a retail survey*. *Tobacco Control*, 27, 684–8.
- John, R., Chelwa, G., Vulovic, V., & Chaloupka, F. (2019). *Using Household Expenditure Surveys for Research in the Economics of Tobacco Control. A Tobacconomics Toolkit*. Chicago, IL: Tobacconomics, Health Policy Center, Institute for Health Research and Policy, University of Illinois at Chicago.
- Joossens, L., Lugo, A., Vecchia, C. L., et ál. (2014). *Illicit cigarettes and hand-rolled tobacco in 18 European countries: a cross-sectional survey*. *Tobacco Control*, 23, e17–23.

Kurti, M., He, Y., von Lampe, K., et ál. (2017). *Identifying counterfeit cigarette packs using ultraviolet irradiation and light microscopy*. *Tobacco Control*, 26, 29–33.

Lakhdar, C. B. (2008). *Quantitative and qualitative estimates of cross-border tobacco shopping and tobacco smuggling in France*. *Tobacco Control*, 17, 12–6.

Liber, A. C., & Warner, K. E. (2018). *Has Underreporting of Cigarette Consumption Changed Over Time? Estimates Derived From US National Health Surveillance Systems Between 1965 and 2015*. *American Journal of Epidemiology*, 187(1), 113–9.

Little, M., Ross, H., Bakhturidze, G., et ál. (2019). *Illicit tobacco trade in Georgia: prevalence and perceptions*. *Tobacco Control*.

Maas, C. J., & Hox, J. J. (2005). *Sufficient sample sizes for multilevel modeling*. *Methodology*, 1, 86–92.

Maldonado, N., Llorente, B. A., Iglesias, R. M., et ál. (2018). *Measuring illicit cigarette trade in Colombia*. *Tobacco Control*.

Merriman, D. (2010). *The Micro-Geography of Tax Avoidance: Evidence from Littered Cigarette Packs in Chicago*. *Am Econ J Econ Policy*, 2, 61–84.

El Ministerio de Finanzas de Polonia. (Consultado en su versión en polaco el 4 de agosto de 2019). *Wzory banderol podatkowych i legalizacyjnych*.

https://www.podatki.gov.pl/media/2993/fsr-win-003-mf-gov-pl-folders_citrix-fqgv-pulpit-wzory-banderol.pdf

Observatorio Chileno de Drogas. (2017). *Décimo Segundo Estudio Nacional de Drogas en Población General*. Santiago de Chile, Chile.

Observatorio Chileno de Drogas. (2015). *Estudios nacionales de drogas en población escolar de Chile 2015*. Instituto Nacional de Estadísticas y Censos. Tercera Encuesta Nacional de Factores de Riesgo para Enfermedades No Transmisibles. Buenos Aires, Argentina.

<http://www.senda.gob.cl/observatorio/estudios/poblacion-escolar/>

Paraje, G., Araya, D., & Drope, J. (2018, Dec 15). *Illicit cigarette trade in Metropolitan Santiago de Chile*. *Tobacco Control*.

Paraje, G. (2018). *Illicit Cigarette Trade in Five South American Countries: A Gap Analysis for Argentina, Brazil, Chile, Colombia and Peru*. *Nicotine & Tobacco Research*.

Patrick, D. L., Cheadle, A., Thompson, D. C., Diehr, P., Koepsell, T., & Kinne, S. (1994). *The validity of self-reported smoking: a review and meta-analysis*. *American Journal of Public Health*, 84(7), 1086–93.

Philip Morris International. (2019) *Información para el inversor, julio de 2019*.

Pizarro, M. E., Giacobone, G., Shammah, C., & Stoklosa, M. (En revisión). *Empty Pack Survey to Estimate Illicit Tobacco Trade in the City of Buenos Aires, Argentina*.

Ross, H. & Blecher, E. (2019). *Illicit Trade in Tobacco Products Need Not Hinder Tobacco Tax Policy Reforms and Increases*. Reporte de *Tobacconomics*. Chicago, IL: *Tobacconomics*, Health Policy Center, Institute for Health Research and Policy, University of Illinois at Chicago.

- Ross, H. (2015). Understanding and measuring tax avoidance and evasion: *A methodological guide*. Washington, D. C. http://www.tobaccoecon.uct.ac.za/sites/default/files/image_tool/images/405/Publications/reports/Understanding-and-measuring-tax-avoidance-and-evasion-A-methodological-guide1.pdf.
- Ross, H., Tesche, J., & Vellios, N. (2017) Undermining Government Tax Policies: *Common strategies employed by the tobacco industry in response to tobacco tax increases*. *Preventive Medicine*, 105, S19–S22
- Ross, H., Vellios, N., Batmunkh, T., et ál. (2019). *Impact of tax increases on illicit cigarette trade in Mongolia*. *Tobacco Control*.
- Rowell, A., Evans-Reeves, K., Gilmore, A. B. (2014). *Tobacco industry manipulation of data on and press coverage of the illicit tobacco trade in the UK*. *Tobacco Control*, 23, e35–43.
- Sejm de la República de Polonia. (2019). *Ustawa z dnia 9 listopada 1995 r. o ochronie zdrowia przed następstwami używania tytoniu i wyrobów tytoniowych*.
- Sejm de la República de Polonia. (Consultado en su versión en polaco el 20 de enero de 2019). *Ustawa z dnia 6 grudnia 2008 r. o podatku akcyzowym. Dz. U. z 2009 r. nr 3, poz. 11. 2008*. <http://prawo.sejm.gov.pl/isap.nsf/download.xsp/WDU20090030011/O/D20090011.pdf>
- StatsSA. (2018). *Mid-year population estimates*. Pretoria: Statistics South Africa.
- Stelmach, R., Fernandes, F. L. A., Carvalho-Pinto, R. M., Athanzio, R. A., Rached, S. Z., Prado, G. F., et ál. (2015). *Comparison between objective measures of smoking and self-reported smoking status in patients with asthma or COPD: are our patients telling us the truth?* *Jornal brasileiro de pneumologia: publicacao oficial da Sociedade Brasileira de Pneumologia e Tisiologia*, 41(2), 124–32.
- Stoklosa, M., & Ross, H. (2014). *Contrasting academic and tobacco industry estimates of illicit cigarette trade: evidence from Warsaw, Poland*. *Tobacco Control*, 23, e35–43.
- Sudman, S. (1983). *Handbook of Survey Research: Applied Sampling*. Nueva York: Academic Press, Inc.
- Szklo, A., Iglesias, R. M., Carvalho de Souza, M., Szklo, M., & Maria de Almeida, L. (2017). *Trends in Illicit Cigarette Use in Brazil Estimated From Legal Sales, 2012–2016*. *American Journal of Public Health*, 108(2), 265–9.
- Naciones Unidas. (Consultado en su versión en inglés el 4 de agosto de 2019). *Protocol to Eliminate Illicit Trade in Tobacco Products*. https://treaties.un.org/pages/ViewDetails.aspx?src=TREATY&mtdsg_no=IX-4-a&chapter=9&lang=en
- Instituto Nacional del Cáncer de Estados Unidos y Organización Mundial de la Salud. (2016). *The Economics of Tobacco and Tobacco Control*. National Cancer Institute Tobacco Control Monograph 21. NIH Publication No. 16-CA-8029A. Bethesda, MD: Departamento de Salud y Servicios Humanos de los Estados Unidos (HHS, por sus siglas en inglés), Institutos Nacionales de la Salud (NIH, por sus siglas en inglés), Instituto Nacional del Cáncer (NCI, por sus siglas en inglés) y Organización Mundial de la Salud.
- Vellios, N., Van Walbeek, C., & Ross, H. (2019) *Illicit cigarette trade in South Africa: 2002 - 2017*. *Tobacco Control*.
- Organización Mundial de la Salud (OMS). (2013) *Protocol to Eliminate Illicit Trade in Tobacco Products*. Ginebra: Organización Mundial de la Salud.

Organización Mundial de la Salud (OMS). (2010) *WHO Technical Manual on Tobacco Tax Administration*. Ginebra: Organización Mundial de la Salud.

Banco Mundial. (2018) *Economics of Tobacco Taxation Toolkit*. Washington: Banco Mundial.

Zubovic, J., Jovanovic, O., Djukic, M., Jolovic, N., Vladislavljevic, M. (venidero). *Study on Tobacco Consumption in Serbia 2019*. Instituto de Ciencias Económicas, Belgrado, Serbia.

Anexos de Metodología y Código

6

Los Anexos de Metodología y Código se basan en los antecedentes de Stata que se analizan en el anexo del conjunto de herramientas anterior. Los comandos individuales de Stata se colocan entre corchetes angulares < > y están en cursiva. Esto es solo para fines ilustrativos. El comando en sí se debe usar sin corchetes para que el código se ejecute en Stata. Los nombres de las variables utilizadas en los ejemplos están en cursiva.

6.1 Selección de puntos de congregación de alto tráfico en Santiago de Chile

El estudio de Santiago de Chile seleccionó puntos de congregación de alto tráfico usando datos sobre el tráfico peatonal en 767 áreas de alto tráfico. Estas áreas de alto tráfico se obtuvieron de la Encuesta de Origen-Destino del Área Metropolitana de Santiago, mientras que los datos de las superficies construidas (en metros cuadrados) por tipo de infraestructura (por ejemplo, edificios comerciales, instalaciones deportivas, escuelas, universidades, edificios de oficinas, instalaciones de salud, parques, etc.) se obtuvieron del registro de superficies construidas de la Dirección General de Impuestos Internos. Estos dos conjuntos de datos se fusionaron para estimar la relación entre el tipo de infraestructura y el tráfico peatonal en la ciudad. Se utilizó el siguiente modelo de mínimos cuadrados ordinarios (MCO):

$$N_i = \beta_1 C_i + \beta_2 S_i + \beta_3 E_i + \beta_4 O_i + \beta_5 Q_i + \beta_6 H_i + \beta_7 T_i + \beta_8 G_i + \epsilon_i$$

donde N_i es el número de visitas al área i (datos de la Encuesta de Origen-Destino), mientras que las variables independientes son las áreas de metros cuadrados de los edificios comerciales (C_i), instalaciones deportivas (S_i), instalaciones educativas y culturales (E_i), edificios de oficinas (O_i), iglesias y lugares culturales (Q_i), instalaciones de salud (H_i), transporte y telecomunicaciones (T_i), y parques (G_i) (datos del registro de superficies construidas). Las estimaciones de los coeficientes de este modelo se utilizaron para predecir el tráfico peatonal en los 66.467 bloques representados en el registro de superficies construidas en Santiago de Chile. Este volumen de tráfico peatonal pronosticado les permitió a los investigadores identificar 424 puntos–bloques de congregación con un tráfico peatonal excepcionalmente alto. Cuarenta de los 424 puntos de congregación fueron seleccionados al azar con probabilidades ponderadas proporcionales al tráfico peatonal en el área. Los puntos de congregación con mayor tráfico tenían más probabilidades de ser seleccionados en la muestra.

6.2 Instrumento de encuesta de muestra

CUESTIONARIO INDIVIDUAL

A1 Ubicación:

A2 Número de sector censal:

A3 Nombre de la calle:

A4 Código del lugar:

A5 Hora de inicio de la entrevista:

A6 Hora de inicio de la entrevista en el sector censal:

A7 Dirección completa:

Acérquese a la casa y pregunte a la primera persona contactada.

Hola, mi nombre es _____ y vengo en nombre de un equipo internacional de investigadores de la salud del Instituto _____. Este estudio está patrocinado por la Sociedad Americana del Cáncer. Estamos realizando una encuesta sobre el tabaquismo. Me gustaría hacerle algunas preguntas a una persona de 18 años o más para ver si alguien de este hogar sería elegible para nuestra investigación, ¿cree que sea posible? Solo le quitaré unos minutos de su tiempo y toda la información será estrictamente confidencial.

B1 ¿Cuántos fumadores de 18 años o más viven en su casa?

|_|_|

88 Se negó a responder

99 No lo sabe

¿Alguno de estos fumadores está en casa en este momento? (Si es así) ¿Puedo hablar con él / ella? (Si es la persona misma) ¿Puedo hablar con usted?

Cuando identifique al fumador que completa la cuota, presente el formulario de consentimiento libre e informado y después comience la entrevista.

Características personales

B2 Sexo:

1 |__| Masculino 2 |__| Femenino

B3 ¿En qué año nació usted? Registrar en 4 dígito

|__| |__| |__| |__|

8888 Se negó a responder

9999 No lo sabe

B4 ¿En qué mes nació usted?

|__| |__|

01 Enero

02 Febrero

03 Marzo

04 Abril

05 Mayo

06 Junio

07 Julio

08 Agosto

09 Septiembre

10 Octubre

11 Noviembre

12 Diciembre

888 Se negó a responder

999 No lo sabe

B5 ¿Qué día nació usted?

|__| |__|

88 Se negó a responder

99 No lo sabe

B6 ¿Cuál es el último grado y nivel de estudios que ha completado?

|_|_|

1.0 Nunca asistió a la escuela regularmente

Escuela primaria

2.0 Jardín de niños

2.1 1.º grado

2.2 2.º grado

2.3 3.º grado

2.4 4.º grado

2.5 5.º grado

2.6 6.º grado

Escuela Secundaria

2.7 7.º grado

2.8 8.º grado

Bachillerato

3.1 9.º grado

3.2 10.º grado

3.3 11.º grado

3.4 12.º grado

4.1 Escuela secundaria técnica o científica, 1 año

4.2 Escuela secundaria técnica o científica, 2 años

4.3 Escuela secundaria técnica o científica, 3 años

Educación superior

5.1 Licenciatura, 1 año

5.2 Licenciatura, 2 años

5.3 Licenciatura, 3 años

5.4 Licenciatura, 4 años

5.5 Licenciatura, 5 años

5.6 Licenciatura, 6 años

5.7 Licenciatura, 7 años

5.8 Licenciatura, 8 años

6.0 Posgrado (especialización, maestría, doctorado)

88 Se negó a responder

TABAQUISMO

Ahora le voy a preguntar sobre su consumo de cigarrillos hechos en fábrica. No incluya cigarrillos de marihuana, cigarrillos armados a mano ni cigarrillos electrónicos en sus respuestas.

B7 Actualmente, ¿fuma algún producto de tabaco?

1. Sí, diariamente
2. Sí, menos que diariamente
8. Se negó a responder
9. No lo sabe

Si seleccionó 8 o 9, complete el cuestionario

Si seleccionó 1 o 2, continúe

B8 En promedio, ¿cuántos cigarrillos fuma por día o por semana?

1. Uno o más, diariamente. ¿Cuántos por día (B9)?
2. Uno o más, semanalmente. ¿Cuántos por semana (B10)?
3. Menos de una vez a la semana. ¿Cuántos por mes (B11)?
4. Menos de uno al mes. ¿Cuántos por año (B12)?
8. Se negó a responder
9. No lo sabe

LOS CIGARRILLOS DE LA ÚLTIMA COMPRA

Las siguientes preguntas son sobre la última vez que compró **cigarrillos hechos en fábrica** para su propio consumo.

B9 La última vez que compró cigarrillos para su propio consumo, ¿cuántos cigarrillos compró?

(Encuestador: Registre la cantidad y, cuando sea necesario, registre los detalles de la unidad)

UNIDAD (20)	CANTIDAD	DETALLE DE LA UNIDAD	
1 <input type="checkbox"/> Cigarrillos	<input type="text"/> (B10) unidades		
2 <input type="checkbox"/> Cajetillas	<input type="text"/> (B11) unidades	¿Cuántos cigarrillos venían en la cajetilla?	<input type="text"/> (B12) cigarrillos
3 <input type="checkbox"/> Cartones	<input type="text"/> (B13) unidades	¿Cuántas cajetillas había dentro del cartón?	<input type="text"/> (B14) cajetillas
4 <input type="checkbox"/> Nunca compré cigarrillos			

B10 ¿Cuánto pagó en total?

\$ |_|_|_|_|_|_|_|_|_|_|_|_|_|_|_|_|

B11 La última vez que compró cigarrillos para su propio uso, ¿qué marca de cigarrillos compró?

_____ |_|_|_|_| (código de la marca) (B17)

B12 ¿Es su marca habitual?

- 1. |_| Sí
- 2. |_| No

Si seleccionó 2, continúe al B19; de lo contrario, continúe al B21.e move on to B21.

B13 ¿Cuál es su marca habitual?

_____ |_|_|_|_| (código de la marca) (B20)

B14 La última vez que compró cigarrillos para usted, ¿dónde los compró?

(Encuestador: espere una respuesta espontánea)

- 1. |_| Tienda de comestibles
- 2. |_| Supermercado
- 3. |_| Vendedor ambulante
- 4. |_| Farmacia
- 5. |_| Kiosco
- 6. |_| *Duty free* (Tienda libre de impuestos)
- 7. |_| Fuera del país
- 8. |_| Internet
- 9. |_| Máquina expendedora (*Vending machine*)
- 10. |_| Otro _____

CARACTERÍSTICAS DE LA CAJETILLA DE CIGARRILLOS

Es muy importante para el estudio conocer y describir algunas características de las cajetillas de cigarrillos que consumen todos los fumadores que participan en él. Me gustaría ver su cajetilla de cigarrillos, anotar sus características y tomarle una fotografía. ¿Tiene una cajetilla de cigarrillos con usted? ¿Puede mostrármela?

(Cuando muestre la cajetilla) ¿Esta es la cajetilla de cigarrillos de su última compra? (Si no es la última cajetilla que compró) ¿Tendrá con usted la última cajetilla que compró? ¿Puede mostrármela? ¿Puedo tomar una fotografía de esta cajetilla?

C1 El informante permitió:

1. Revisión y fotografía de la cajetilla
2. Solo revisión de la cajetilla
3. Solo fotografía de la cajetilla

C2 ¿Es la cajetilla vista/fotografiada la última que compró?

1. Sí
2. No

C3 Código de la marca:

999: Marca de cigarrillos no registrada con las autoridades nacionales

Si seleccionó 999 continúe al C4, de lo contrario pase al C5

C4 Nombre de la marca: _____

C5 ¿Hay un sello fiscal oficial visible en la cajetilla?

1. Sí, la cajetilla tiene un sello oficial visible
2. La cajetilla tiene un sello visible, pero este sello es de otro país
3. Tiene fragmentos del sello fiscal y el origen del sello es difícil de definir
4. No, no hay sello y no se ve ningún fragmento del sello.

C6 ¿Hay imágenes de advertencia sanitaria en la cajetilla?

1. Sí
2. No

Si seleccionó 1, continúe al C7

Si seleccionó 2, pase al C10

C7 ¿Cuál es el idioma del mensaje de texto que acompaña a la imagen de advertencia?

1. Inglés
2. Español
3. Portugués
4. Otro

C8 Código de la imagen de advertencia:

99 Una imagen diferente.

C9 Tamaño de la advertencia (imagen + texto) en relación con el lado en el que se expone?

1. Menos del 30 %
2. 30%
3. Mayor al 30 %

Marque si alguno de los siguientes artículos está presente en la cajetilla:

C10 Light, ultralite

1. Sí
2. No

C11 Nombre del país al que se dirige la cajetilla:

1. Sí Si es así, qué país _____
2. No

6.3 Muestra de solicitud de exención presentada a el Consejo de Revisión Institucional

Preguntas de selección para la exención

Si responde 'Sí' a cualquiera de las preguntas de la A la D que aparecen a continuación, entonces DETÉNGASE y utilice el formulario de solicitud para la revisión inicial de la IRB.

Si contesta 'No' a todas las preguntas de la A la D que aparecen a continuación, continúe llenando esta solicitud de exención.

Importante: Incluya un formulario de evaluación llenado con su solicitud

A. Para las investigaciones que involucran poblaciones, intervenciones o manipulaciones especiales

1. ¿Su investigación involucra a mujeres embarazadas, fetos o prisioneros? Sí No
2. ¿Su investigación incluye el uso de procedimientos de encuestas o entrevistas con niños? Sí No
3. ¿Su investigación incluye la observación de niños en entornos en los que los investigadores participarán en las actividades que se están observando? Sí No

B. Para las investigaciones que utilicen procedimientos de encuesta, procedimientos de entrevista, procedimientos de observación y cuestionarios

1. Si los datos se van a grabar en una cinta de audio o de vídeo, ¿existe un daño potencial¹ para los sujetos si la información se revela o se divulga? Sí No
2. Si los sujetos deben ser identificables por su nombre o por sus datos demográficos, ¿existe un daño potencial para los participantes si se revela la información? Sí No
3. ¿La recopilación de datos incluirá información sensible (por ejemplo, actividades ilegales o temas sensible como la orientación sexual, conducta sexual, conducta laboral indeseable u otros datos que puedan ser dolorosos o muy vergonzosos de revelar, como la muerte de un familiar, recuerdos de abusos físicos)? Sí No

C. Para realizar investigación que utiliza datos, documentos, registros o especímenes existentes² o archivados únicamente

1. ¿Se recopilarán datos, documentos, registros o especímenes de los sujetos después de la presentación de esta solicitud?
2. Si los datos, documentos, registros o especímenes están etiquetados originalmente de tal manera que los sujetos puedan ser identificados, directa o indirectamente a través de enlaces de identificación, ¿el investigador está registrando los datos para los fines de esta investigación de tal manera que los sujetos puedan ser identificados, directa o indirectamente a través de enlaces de identificación (es decir, información demográfica que pueda razonablemente conducir a la identificación de los sujetos individuales: nombre, número de teléfono; o cualquier número de código que pueda utilizarse para vincular los datos del investigador con el registro de origen; número de registro médico o número de admisión al hospital)? Sí No

¹ Daño a los sujetos significa que cualquier divulgación de las respuestas de los sujetos humanos fuera de la investigación podría poner razonablemente a los sujetos en riesgo de responsabilidad penal o civil o podría ser perjudicial para la posición financiera, empleabilidad o reputación de los sujetos.

² Existentes significa que los artículos existen antes de que se propusiera la investigación o se recogieron antes de la investigación para un propósito distinto al de la investigación propuesta.

D. Para realizar investigación que utiliza información de salud protegida

1. ¿La investigación implicará el uso o la divulgación de información de salud identificable individualmente que incluya: nombres, fechas (que no sean años), números de teléfono, números de fax, direcciones electrónicas de correo electrónico, números de seguridad social, números de registros médicos, números de beneficiarios de planes de salud, números de cuenta, números de certificados/licencias, identificadores de dispositivos y números de serie, direcciones URL, direcciones de Internet, identificadores biométricos, cara completa o imágenes comparables, o cualquier número, característica o código de identificación único?
- Sí No

6.4 Muestreo de regiones geográficas utilizando el método de probabilidad proporcional al tamaño (PPT) de prevalencia de tabaquismo prevista en Brasil

Los investigadores del estudio brasileño comenzaron su muestreo de la región geográfica investigando el número de fumadores que viven en cada zona censal. Su recolección de cajetillas desechadas se implementó con base en estos resultados. Los investigadores utilizaron datos de la Encuesta Nacional de Salud para estimar la probabilidad de fumar según las características socioeconómicas. Estos resultados dieron a los investigadores el número aproximado de fumadores en cada zona censal.

El código de Stata utilizado es el siguiente. El procedimiento comienza declarando el diseño de la Encuesta Nacional de Salud. Identifica la variable principal de la unidad de muestreo del estudio y la variable de ponderación. Aquí, los nombres de las variables eran `psu` y `weight`, respectivamente.

```
svyset psu [pweight=weight]
```

Las variables incluidas en el modelo son:

- *fumprodind*: uso de cigarrillos manufacturados (0: no, 1: sí); esta es la variable dependiente en el modelo
- *idadetest*: grupo de edad (0: 18–24, 1: 25–34, 2: 35–44, 3: 45–54, 4: 55–64, 5: 65+)
- *sexo*: sexo (0: masculino, 1: femenino)
- *renda_total*: ingreso mensual promedio del propietario del hogar

Se utilizó la siguiente regresión logística para estimar la relación entre el tabaquismo y las características socioeconómicas:

```
xi:svy:logit fumprodind i.sexo i.idadetest renda_total
```

donde `<xi>` permite la inclusión de variables categóricas (por ejemplo, la edad) en el modelo, mientras que `<svy>` ajusta los resultados de las ponderaciones de la encuesta previamente definidas. El modelo utiliza variables de edad categorizadas. Esta es una buena práctica, porque la relación entre fumar y la edad no suele ser lineal. Si los ingresos no son una variable disponible ni en la encuesta ni en los datos individuales de la UPM, otras variables, como la educación, pueden servir de aproximación a los ingresos.

Los resultados estimados del modelo se pueden utilizar para predecir la prevalencia del tabaquismo en cada zona censal utilizando el comando `<argins>` del programa Stata y las proporciones de la población adulta con cada característica dentro de cada zona censal:

```
argins, at(_Isexo_1="VALUE=PROPORTION OF FEMALES"  
_Iidadetest_1="VALUE=PROPORTION OF 25-34"  
_Iidadetest_2="VALUE=PROPORTION OF 35-44"  
_Iidadetest_3="VALUE=PROPORTION OF 45-54"  
_Iidadetest_4="VALUE=PROPORTION OF 55-64"  
_Iidadetest_5="VALUE=PROPORTION OF 65+"  
renda_total = "VALUE= AVERAGE MONTHLY INCOME OF THE HOUSEHOLD  
OWNER")
```

Supongamos que el 60% de la población es femenina, basado en la zona censal distribuida de tal manera que el 32% tiene entre 25 y 34 años, el 23% tiene entre 35 y 44 años, el 10% tiene entre 45 y 54 años, el 5% tiene entre 55 y 64 años y el 4% tiene 65 años o más. El ingreso mensual promedio del propietario del hogar es de 3.800 reales brasileños. El comando para predecir la prevalencia del tabaquismo en esa zona censal es:

```
argins, at(_Isexo_1=0.6 _Iidadetest_1=0.32 _Iidadetest_2=0.23 _Iidadetest_3=0.1  
_Iidadetest_4=0.05 _Iidadetest_5=0.04 renda_total=3800)
```

El número de fumadores en cada zona censal es el producto de la prevalencia estimada de fumadores adultos en cada UPM y el número total de adultos en la UPM.

La etapa final del proceso de muestreo requiere la extracción aleatoria de las UPM. La probabilidad de ser seleccionado es proporcional al número de fumadores de cada UPM en relación con el total. La función de Stata `<gsample>` se utiliza para el muestreo con probabilidades proporcionales al tamaño (PPT). Si la función aún no está instalada, se puede utilizar el siguiente comando:

```
ssc install gsample
```

La función `<gsample>` requerirá el paquete `<moremata>` para ejecutarse. El siguiente comando instalará el paquete:

```
ssc install moremata
```

El conjunto de datos de la UPM debe estar organizado en formato largo; cada fila es una UPM diferente. La variable `'pop'` representa el número de fumadores en cada UPM. La variable de ponderación de la muestra se crea con los siguientes comandos:

```
egen popsum=sum(pop)  
gen sweight= (pop/popsum)
```

La función `<gsample>` y las ponderaciones calculadas previamente se utilizan para extraer una muestra de n UPM sin reemplazo. En el ejemplo siguiente, $n = 50$.

```
gsample 50 [aw=sweight], wor
```

6.5 Muestreo aleatorio simple

A diferencia de PPT presentado en el ejemplo anterior, Stata no requiere un nuevo procedimiento para ejecutar un muestreo aleatorio simple. Supongamos que se proporciona una lista completa de UPM en una ciudad determinada. Si los investigadores necesitan seleccionar el 5% de las UPM para la muestra, el comando de Stata a continuación dará el 5% de la muestra:

```
sample 5
```

Si la lista contiene las UPM de diferentes ciudades, con las ciudades identificadas por la variable *city*, entonces los investigadores pueden dibujar una muestra aleatoria simple del 5% de las UPM por ciudad usando el siguiente comando:

```
sort city  
by city: sample 5
```

6.6 Ponderación multivariante: ejemplo de Chicago

Supongamos que, al igual que Merriman (2010), los investigadores quieren dibujar una muestra aleatoria ponderada de regiones geográficas. Las zonas con una población más numerosa (ponderación del 100%) y niveles de empleo más elevados (ponderación del 50%) tienen más posibilidades de ser seleccionadas en la muestra. La idea es imitar el comportamiento del fumador promedio, de tal manera que dos tercios de los cigarrillos se desechan en casa y un tercio en el trabajo.

En primer lugar, es necesario crear una base de datos que contenga todo el universo de regiones geográficas. La base de datos debe contener las siguientes variables:

- *id*: número de id de una región geográfica determinada
- *pop*: población de fumadores que viven en esa región geográfica
- *emp*: población de fumadores que trabajan en esa región geográfica

Los siguientes comandos se pueden utilizar para generar ponderaciones de muestreo:

```
egen popsum=sum(pop)  
egen empsum=sum(emp)  
gen weight= (pop/popsum)+0.5*(emp/empsum)
```

La función `<gsample>` y las ponderaciones de la encuesta se pueden utilizar para extraer una muestra de *n* regiones geográficas sin reemplazo. En el ejemplo siguiente, *n*=23:

```
gsample 23 [aw=weight], wor
```

6.7 Imputar el estado de cumplimiento para fumadores que no mostraron cajetillas a los encuestadores: enfoque de regresión logística

En el conjunto de datos creado, la variable *compl* indica si la cajetilla mostrada por el fumador tenía impuestos pagados o no. Faltan los valores de la variable para los fumadores que no mostraron sus cajetillas. Los investigadores determinaron que las siguientes variables son buenos predictores de la variable *compl*. La regresión logística de *compl* en esas variables produce altos valores de Pseudo-R² y bajos valores de AIC y BIC.

- *compl*: cigarrillos que cumplen con las normas (con impuestos pagados) (0: no, 1: si)
- *age*: grupo de edad (0: 18–24, 1: 25–34, 2: 35–44, 3: 45–54, 4: 55–64, 5: 65+)
- *sex*: sexo (0: masculino, 1: femenino)
- *brand*: marca de cigarrillos (valores 0 a 10)

La probabilidad de cumplimiento de los impuestos para los fumadores que no mostraron sus cajetillas puede predecirse mediante:

```
xi: logit compl i.age i.sex i.brand  
predict pred, p
```

Para asignar el cumplimiento o el incumplimiento de los fumadores con los datos de *compl* que faltan, debe elegirse un punto de corte para la variable *pred*. El punto de corte estándar es 0,5:

```
gen compl2 = 0  
replace compl2=1 if pred>0.5
```

Se puede comparar el cumplimiento real e imputado de las observaciones con los valores de *compl* no faltantes:

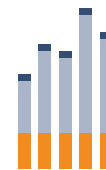
```
tab compl compl2
```

Si es necesario, se pueden elegir otros puntos de corte. Por último, los valores imputados de cumplimiento deben asignarse a las observaciones con valores de *compl* que faltan:

```
replace compl=compl2 if compl==.
```



INSTITUTE FOR
HEALTH RESEARCH
AND POLICY



tobacconomics

Economic Research Informing
Tobacco Control Policy

www.tobacconomics.org

[@tobacconomics](https://twitter.com/tobacconomics)

