



Centro de Investigación en Matemáticas, AC

“Evasión por Subfacturación y Subdeclaración”

Investigadores

**Graciela Gonzáles Farías
Víctor Muñiz Sánchez
Edgar Jiménez Peña
Héctor Hernández González
José Ramón Domínguez Molina**

Contenido

Resumen ejecutivo.....	1
Introducción.....	3
1. Antecedentes.....	5
2. Descripción de datos	7
Estadística Descriptiva Personas Morales.....	9
Estadística Descriptiva Personas Físicas	10
3. Metodología: Caracterización de la Sub-facturación y Sub-declaración.....	11
Caso Personas Morales.....	13
Caso Personas Físicas	14
Detección de Evasión	16
Análisis Personas Morales	17
Análisis Personas Físicas.....	20
4. Análisis de Evasión	23
Caso Personas Morales.....	24
Caso Personas Físicas	28
5. Conclusiones y recomendaciones	33
Referencias	35
Anexo.....	36
Muestra de análisis.....	36
Bases de datos.....	42
Estandarización de la información	44
Cálculo de ISR	47
Tasas de evasión de contribuyentes que evaden impuestos.....	48
Análisis de aprendizaje automático.....	51

Resumen ejecutivo

El presente estudio tiene como objetivo calcular el potencial de recaudación de los contribuyentes que en su actividad económica suelen emitir tickets, esto con el fin de estimar la evasión por sub-declaración de ingresos. Este análisis es uno de los primeros que relaciona las facturas electrónicas (CFDI) emitidas y recibidas con las declaraciones y pagos mensuales a gran escala mediante el procesamiento de información de cerca de mil millones de facturas electrónicas emitidas por personas morales del régimen general y por personas físicas con actividad empresarial.

La importancia de esta investigación radica en el manejo de datos con un gran nivel de detalle de todas las facturas emitidas y recibidas por un conjunto de contribuyentes correspondiente a 2015 y 2016. En este sentido, se utilizó el análisis descriptivo y de aprendizaje automático para determinar las características más importantes de los contribuyentes que tienen indicios de generar sub-declaración.

Entre las acciones que se realizaron para un correcto análisis de datos se llevó a cabo la estandarización, homogenización, la organización, adecuación y limpieza de la información de los CFDI para detectar, entre otros defectos, facturas atípicas, montos extraordinarios y omisión de campos.

En cuanto a los contribuyentes analizados se tomó a aquellos pertenecientes a sectores cuya actividad económica les permite generar tickets. En suma, se seleccionaron 7 sectores de la economía compuestos por 144 actividades económicas.

La metodología consistió en comparar la sumatoria de los montos de ventas facturados a nivel contribuyente-mes contra la declaración de ingresos. Para dicho comparativo se utilizaron bases de datos provenientes del SAT tanto de CFDI como de Declaraciones y Pagos. Adicionalmente, se define el cociente de subdeclaración como la razón de los ingresos por facturas, definido como la sumatoria de ingresos facturados antes de impuestos indirectos entre el ingreso nominal que declara en el mes:

$$\text{Cociente de subdeclaración} = \frac{\text{Ingreso por facturas}}{\text{Ingreso nominal que declara en el mes}}$$

El cual indica que si:

- El valor es mayor a uno, existe más ingreso por facturas que el declarado, por lo tanto, es factible que exista sub-declaración de ingresos.
- El valor es igual a la unidad, los contribuyentes declaran al final de mes lo mismos montos de sus facturas emitidas.
- El cociente es menor a uno, existe menos ingreso por facturas que lo declarado al final del mes, por lo tanto, es factible que exista un comportamiento de sub-facturación. Para fines

del estudio, este caso, no se consideró como evasión, ya que existen ingresos no facturados como la ganancia cambiaria, los rendimientos, el interés acumulable, entre otros.

Se determinó mediante criterios estadísticos que si el cociente de sub-declaración se encuentra entre 0.9206 y 1.0792 para personas morales y entre 0.8123 y 1.1531 para personas físicas no presentan indicios de que el contribuyente esté sub-facturando o sub-declarando, pues existe la posibilidad de errores de registro o comportamientos contables que permitan ciertas variaciones alrededor de la unidad.

Con respecto a los resultados, para las personas morales, es posible que existan ingresos no declarados de alrededor del 12.0% para 2015 y 12.6% para 2016, esto representa una evasión sobre el potencial de recaudación de 10.4% para 2015 y 10.2% para 2016 (los contribuyentes que fueron identificados sub-declarando ingresos presentaron una evasión promedio de 37.1% y 34.0% para 2015 y 2016 respectivamente).

Par el caso de las personas físicas, es posible que existan ingresos no declarados de alrededor del 9.9% para 2015 y 9.8% para 2016, esto representa una evasión sobre el potencial de recaudación de 67.4% para 2015 y 57.7% para 2016 (los contribuyentes que fueron con sub-declarando ingresos presentaron una evasión promedio de 93.4% y 89.5% para 2015 y 2016 respectivamente).

Introducción

El Código Fiscal de la Federación, en el artículo 29 menciona que cuando las leyes fiscales lo establezcan, los contribuyentes tendrán la obligación de expedir Comprobantes Fiscales Digitales por Internet (CFDI), por los ingresos que se perciban o por las retenciones de contribuciones que efectúen. La expedición de CFDI podrá realizarse a través de un Proveedor Autorizado de Certificación (PAC) o con las herramientas disponibles en la página de internet del SAT.

Por otro lado, el marco legal fiscal permite a los contribuyentes, que enajenan bienes o prestan servicios al público en general, expedir una factura electrónica diaria, semanal o mensual donde consten los importes totales correspondientes a los comprobantes de operaciones realizadas con el público en general. Por este motivo los contribuyentes tienen la posibilidad de emitir tickets sin ninguna validez fiscal.

Ésta facilita la supresión de ventas, pues el contribuyente puede no registrar algunas operaciones en efectivo por las que sólo se solicitó tickets. Para contrarrestar estos actos se requeriría que la autoridad fiscal auditara a cada uno de los contribuyentes lo cual es inviable asumiendo que las administraciones tributarias operan con recursos escasos.

Por lo tanto, para combatir la supresión de ventas o también llamada sub-declaración de ingresos, se han incorporado medidas tecnológicas que han dado a las administraciones tributarias la oportunidad de prevenir y detectar este tipo de conductas, con una mejora significativa en la recaudación y en el incremento de eficiencia en sus operaciones.

Según la OCDE (2017) algunos países han incorporado a su sistema de facturación electrónica cajas registradoras que permiten que los tickets sean a la vez comprobantes fiscales, los cuales deben contar con características que permitan asegurar la integridad de los datos y la identidad del emisor.

En México la incorporación del CFDI está llevando al Servicio de Administración Tributaria a realizar auditorías electrónicas que permitan conocer si el contribuyente está cumpliendo correctamente con sus obligaciones.

Este documento es uno de los primeros análisis basado en la información de millones de CFDI en el cual se hace un cruce con las declaraciones de los contribuyentes a partir de algoritmos capaces de detectar patrones o conductas cuando los contribuyentes omiten los registros de ingreso o de facturas emitidas, en otras palabras cuando sub-facturan o sub-declaran, esto para identificar la recaudación potencial según el nivel de actividad económica del contribuyente y poder cuantificar el monto de evasión cuando se declara más o menos del monto facturado, así como, conocer cuáles son los contribuyentes que presentan este comportamiento.

Con el desarrollo de un modelo de predicción de probabilidad de incidencia en eventos de sub-facturación y sub-declaración con base en metodologías de aprendizaje de máquina se pretenden lograr estimaciones confiables de la evasión generada por la sub-declaración, así como poder

generar una serie de indicadores que permitan identificar comportamientos de evasión por sub-declaración con suficiente confianza estadística. Así como cuantificar la sub-facturación.

El CIMAT tiene la oportunidad de ofrecer al SAT y a la sociedad desarrollos científicos con base en su amplia experiencia en las áreas de “Estadística Aplicada”, “Data Analytics” y “Big Data”, bajo la consigna de un aprovechamiento máximo de la información y la utilización de metodologías de vanguardia para la construcción de los algoritmos de aprendizaje.

El presente documento se compone en su primera sección por los antecedentes relevantes para el estudio, en la segunda sección se tiene una descripción de las bases de datos involucradas así como su estandarización y homogenización, en la tercera sección se plantea la metodología de análisis para el objetivo principal de este estudio, en la cuarta sección se tiene el análisis de evasión por sub-declaración de personas morales y físicas en general, por mes, así como por entidad federativa y sector de la economía. Por último, se tiene una sección de conclusiones, referencias y anexos técnicos.

1. Antecedentes

Una de las principales fuentes de recursos para los estados es la recaudación de impuestos, hecho que para México se establece en el artículo 31 de la Carta Magna el cual instituye la obligación de los mexicanos para contribuir en el gasto público, de manera proporcional y equitativa que dispongan las leyes. No cumplir con las disposiciones legales referentes a las contribuciones coloca a los contribuyentes en una situación de evasión fiscal¹.

La evasión fiscal, según la OCDE (2017), no solo merma los recursos públicos, sino también afecta la competitividad, ya que pone en desventaja a los contribuyentes cumplidos haciendo que estos tengan gastos superiores a los que no pagan impuestos.

Según el estudio de la OCDE (*Ibidem*), una de las principales formas de evasión que se da alrededor del mundo es la llamada sub-declaración, que consiste en declarar menores ingresos mediante la supresión de ventas registradas, así mismo, se menciona que la sub-declaración de ingresos se encuentra generalmente en negocios con venta directa al consumidor final, como restaurantes, bares, taxis, tiendas de conveniencia; entre otras.

Detectar este tipo de evasión en el pasado era muy costoso y le llevaba mucho tiempo a las administraciones tributarias; sin embargo, el uso de la tecnología ha cambiado este escenario, por ejemplo, la obligatoriedad de emitir Certificados Fiscales Digitales por Internet (CFDI) ha permitido a las autoridades detectar comportamientos evasores de una manera más eficiente y eficaz.

Para la prevención y detección de delitos fiscales las autoridades al redor del mundo han utilizado dispositivos electrónicos que garantizan la integridad de los datos, por lo que, al tener acceso vía online la administración fiscal cuenta con una transmisión automática y periódica de los datos que permite tener acceso a la información con propósitos de auditoria.

Por ejemplo, la regulación y certificación de cajas registradoras, asegura que no pueda modificarse su información a través de diversos programas o dispositivos ya que el almacenamiento de su información se encuentra tanto dispositivos internos como externos. Además, la firma digital y las especificaciones que deben contener los tickets de compra o recibos permite tener un control de éstos.

El uso de herramientas tecnológicas no sólo beneficia al fisco, también hay beneficios para las empresas, como lo son:

¹ La evasión de impuestos se da cuando el contribuyente no paga la contribución después de haber incurrido en una situación que la ley marca como objeto gravado (hecho imponible o como hecho generador de la obligación fiscal). La evasión fiscal se da cuando el contribuyente realiza las situaciones jurídicas o de hecho previstas en la Ley del Impuesto Sobre la Renta, Ley del Impuesto al Valor Agregado, entre otras (art. 6 del Código Fiscal de la Federación), que dan origen a las contribuciones, y éste no las paga.

- 1) Garantizar la seguridad de los datos protegiendo a los empresarios de robos por parte de sus empleados.
- 2) Reduce la carga de una auditoría, favoreciendo tanto al contribuyente como a las autoridades.

Así mismo, con la ayuda de estos avances tecnológicos, también se puede estudiar el fenómeno de la subfacturación, el cual se caracteriza por declarar más ingresos de lo registrado en ventas. Este último concepto no puede ser denominado como evasión pues los contribuyentes pueden tener otros ingresos ajenos a su actividad principal (enajenación de bienes o prestación de servicios), tales como ingresos nominales por la ganancia cambiaria, los rendimientos, el interés acumulable, efectos inflacionarios favorables, entre otros.

2. Descripción de datos

Los datos provienen de tres fuentes de información, por un lado, las bases que contienen la información de los CFDI's por contribuyente, por otro, las bases que contienen la información de declaraciones y pagos, por último, las bases que vinculan a cada contribuyente con sus elementos categóricos tales como la entidad federativa y sector económico.

La información proporcionada por el SAT comprende los años 2015 y 2016 la cual está compuesta por datos con periodicidad mensual para la información proveniente de la base de Declaraciones y Pagos (DyP), mientras que la información de CFDI tiene una periodicidad continua porque fue generada en tiempo real. Con el fin de homogenizar la información los análisis se efectuarán en periodos mensuales.

Para fines de este estudio se seleccionó un padrón de contribuyentes compuesto por personas morales de régimen general que por sus actividades tienden a emitir tickets (o realizan ventas en mostrador), que son el tipo de negocio que se inclinan a sub-declarar ingresos según estudios de la OCDE (op. cit.).

Dicha elección se basó en el Sistema de Clasificación Industrial de América del Norte² (SCIAN), el cual consta de cinco niveles de agregación: sector, subsector, rama, sub-rama y clase de actividad económica, por lo que se seleccionaron las actividades económicas que utilizan cajas registradoras en sus actividades específicas.

De lo anterior, se extrajeron 7 sectores, de los cuáles se seleccionaron 144 clases o actividades económicas, por considerar que usan caja registradora y, por tanto, emiten tickets (véase anexo 1).

Sectores SCIAN contemplados	
Sector	Número de actividades
46. Comercio al por menor	63
53. Servicios inmobiliarios y de alquiler de bienes muebles e intangibles	12
48-49. Transportes, correos y almacenamiento	10
62. Servicios de salud y asistencia social	3
71. Servicios de esparcimiento culturales y deportivos, y recreativos	18
72. Servicio de alojamiento y preparación de alimentos	10
81. Otros servicios excepto actividades gubernamentales	28
Total	144

Fuente. Elaboración con datos del SAT.

² Cada país tiene su propia versión del SCIAN, pero se reconoce el nivel en base a un acuerdo trilateral de los países que conforman este sistema (Canadá, Estados Unidos y México). El sector, es el nivel más general; la clase, el más desagregado. El sector se divide en subsectores. Cada subsector está formado por ramas de actividad, las cuales se dividen a su vez en subramas. Las clases, por su parte, son desgloses de las subramas

Por ejemplo, en el sector de comercio al por menor se consideran a las actividades que incluyen el comercio de supermercados, minisuper, tiendas departamentales, tiendas especializadas, entre otros. Con respecto a los servicios de esparcimiento, culturales, deportivos y recreativos, se incluyen los boliches, parques recreativos, zoológicos, museos, clubes, entre otros. Así mismo, para los servicios de alojamiento y de preparación de alimentos se encuentran los restaurantes, bares, cantinas, entre otros.

Dentro de las 144 actividades seleccionadas se encontraron 66,255 personas morales y se seleccionaron los contribuyentes que presentan ingresos tributarios brutos entre 2 millones y 64 millones de pesos³, por lo que la selección de personas morales se reduce a 4,568 contribuyentes. Para el caso de personas físicas se seleccionó la población completa correspondiente a estos sectores, cerca de 100 mil contribuyentes.

En el siguiente cuadro se presentan el número de facturas emitidas por los contribuyentes seleccionados.

Facturas emitidas por tipo de contribuyente		
Persona	2015	2016
Morales	210,646,980	217,602,351
Físicas	178,823,163	186,235,252
Total	389,470,143	403,837,603

Fuente. Elaboración con datos del SAT.

Se consolida la información de las 3 fuentes de información antes descritas y se genera una base de datos compuesta por los elementos categóricos como: el RFC (o identificador asignado), ejercicio (2015 ó 2016), mes, entidad federativa, sector económico y las variables cuantitativas como Ingresos declarados e ingresos percibidos según la información del CFDI. A modo de ejemplificar la información se presenta la siguiente tabla:

ID_RFC	Año Fiscal	Mes que Declara	Entidad Federativa	Sector de la Economía	Ingreso Nominal en el Mes de Declaraciones y Pagos	Ingreso por Facturas Emitidas
3406691	2015	Enero	SINALOA	46	20,799,558	36,057,065
3423745	2015	Enero	MEXICO	81	11,001,060	21,872,414
3423854	2015	Enero	COLIMA	46	28,896,034	49,989,187
3424715	2015	Enero	SAN LUIS POTOSI	46	6,505,046	7,235,791

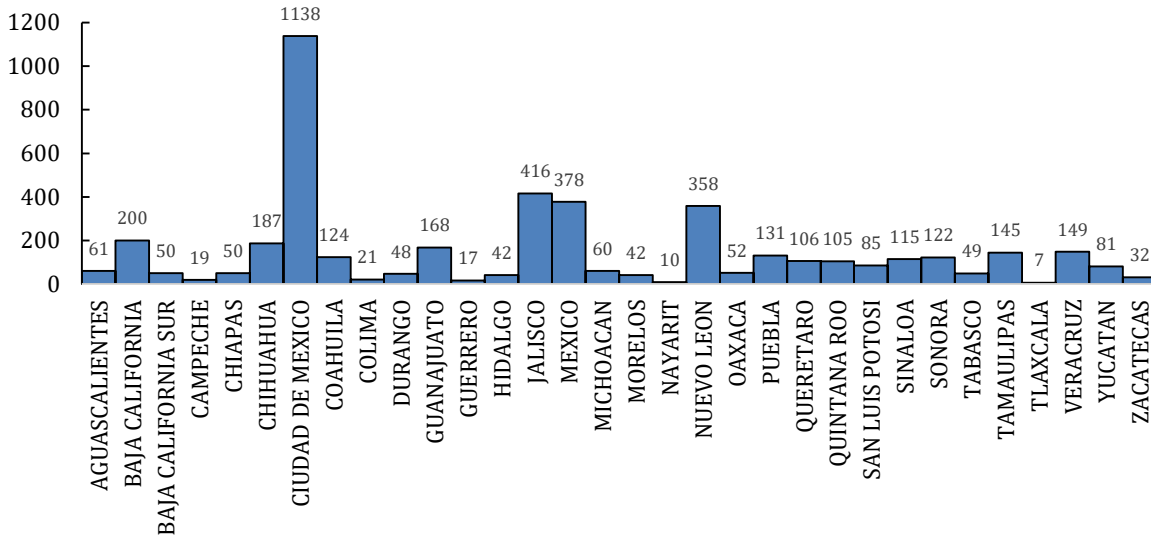
Fuente. Elaboración con datos del SAT.

³ Se optó por contribuyentes de tamaño medio, debido a que según evidencia anecdótica las empresas de gran tamaño suelen usar estrategias fiscales en vez de sub-declarar ingresos.

Estadística Descriptiva Personas Morales

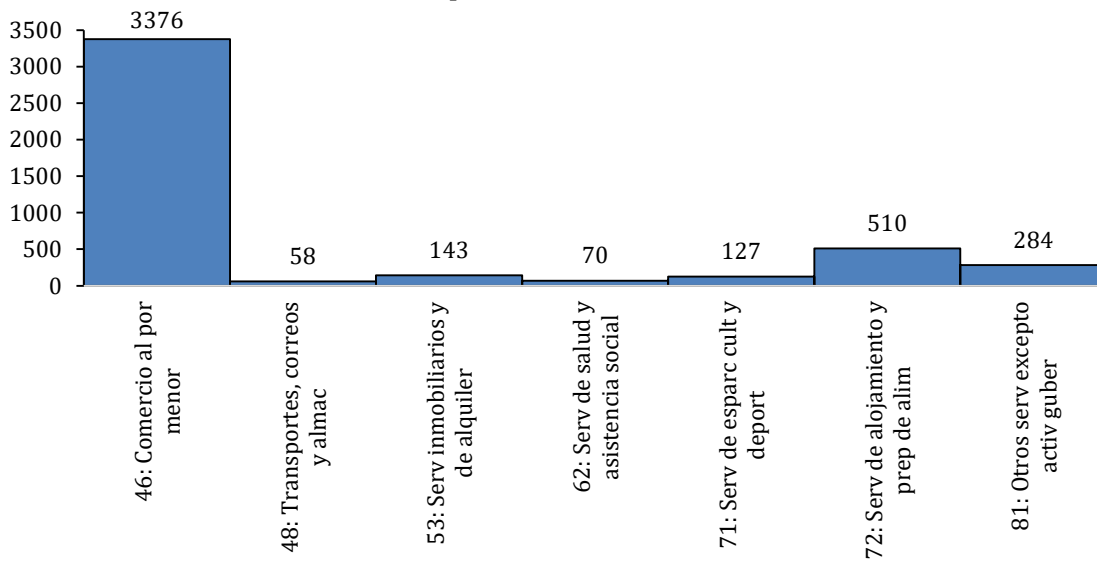
Las bases de datos homogenizadas constan de 4,568 contribuyentes que representan el 6.9% de las personas morales de las actividades seleccionadas, las cuales están distribuidas por estado y sector. Las mayores densidades de contribuyentes se encuentran registrados en la Ciudad de México, Jalisco, Nuevo León y Estado de México, en conjunto representan 41.9% de las personas morales analizadas.

Cantidad de contribuyentes por entidad federativa



Fuente. Elaboración con datos del SAT.

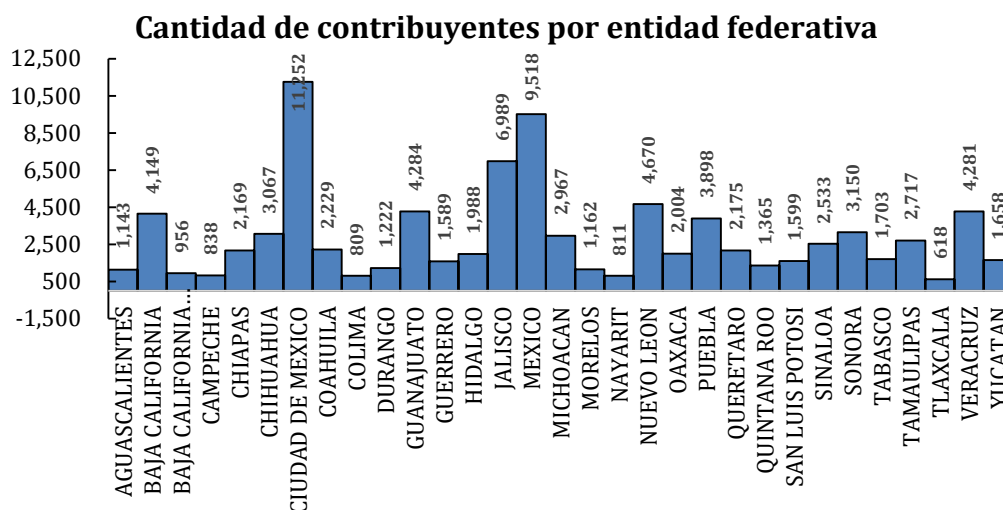
Como era de esperarse la mayor densidad de contribuyentes con ventas en mostrador que emiten tickets están en el sector de comercio al por menor con 73.9%.



Fuente. Elaboración con datos del SAT.

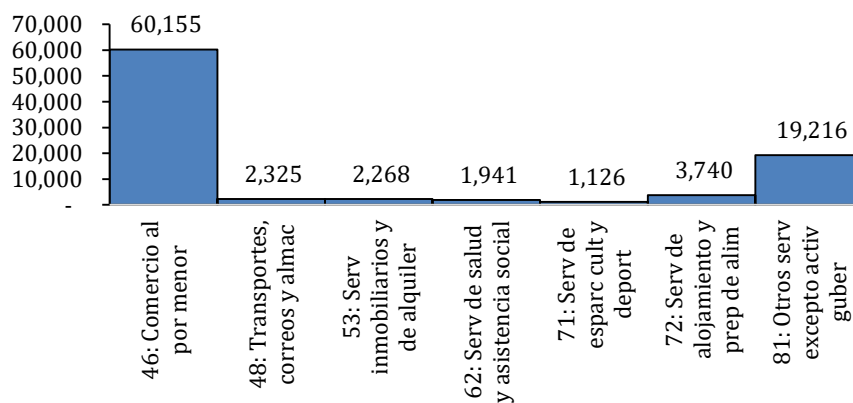
Estadística Descriptiva Personas Físicas

Las bases de datos homogenizadas constan de 90,771 personas físicas de las actividades seleccionadas, las cuales están distribuidas por estado y sector. Las mayores densidades de contribuyentes se encuentran registrados en la Ciudad de México, Jalisco, Nuevo León y Estado de México, en conjunto representan 35.7% de las personas físicas analizadas.



Fuente. Elaboración con datos del SAT.

Al igual que las personas morales, la mayor densidad de contribuyentes con ventas en mostrador que emiten tickets están en el sector de comercio al por menor con 66.3%.



Fuente. Elaboración con datos del SAT.

3. Metodología: Caracterización de la Sub-facturación y Sub-declaración

Como base primordial de la metodología se comparó la sumatoria de los montos de ventas facturados provenientes de la base de CFDI a nivel contribuyente por mes, contra la declaración de ingresos proveniente de la base denominada “Declaraciones y Pagos” (ingresos declarados). Para dicho comparativo se define el cociente de sub-declaración de la siguiente manera:

$$\text{Cociente de subdeclaración} = \frac{\text{Ingreso por facturas}}{\text{Ingreso nominal que declara en el mes}}$$

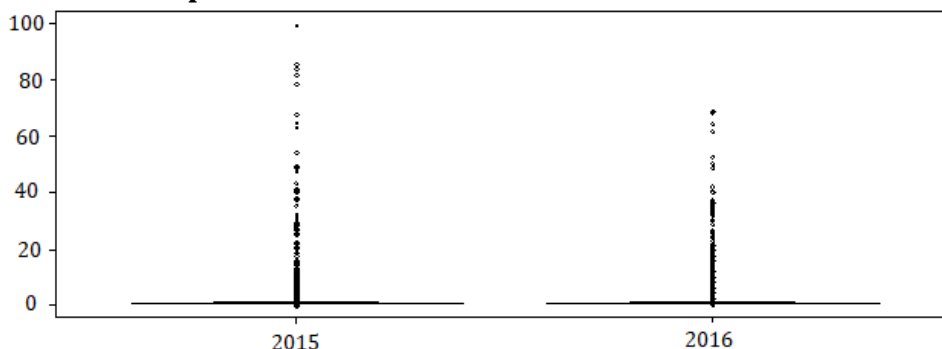
El cual indica que si:

- El valor es mayor a uno, existe más ingreso por facturas que lo que se declara, por lo tanto, es factible que exista sub-declaración de ingresos. Adicionalmente, un supuesto toral para el análisis es asumir que todos los CFDI’s emitidos que son estadísticamente diferentes a la unidad, no son generados por ingresos que provienen de ventas ajenas a la actividad económica del contribuyente.
- El valor es igual a la unidad, los contribuyentes declaran al final de mes la misma cantidad de las facturas que emiten.
- El cociente es menor a uno, existen menos ingresos por facturas que lo que se declara al final del mes, por lo tanto, es factible que exista un comportamiento de sub-facturación. Este caso para el estudio no se consideró como evasión, ya que existen ingresos no facturados como la ganancia cambiaria, los rendimientos, el interés acumulable, venta de activos, entre otros.

A partir de la información resultante del cociente, se generaron gráficas de cajas y brazos (boxplot) para ambos periodos. Se observa que la mediana, sin tomar en cuenta las observaciones atípicas (se tomó como observación atípica aquella que era mayor a la media más 5 desviaciones estándar), es de 1.0 en ambos años, lo que implica que los contribuyentes que se encuentran en medio de la distribución no presentan sub-declaración o algún comportamiento de sub-facturación.

Al realizar la gráfica con todos los datos, se observa que en ambos años se tiene un sesgo importante del cociente de sub-declaración al mostrar valores altos.

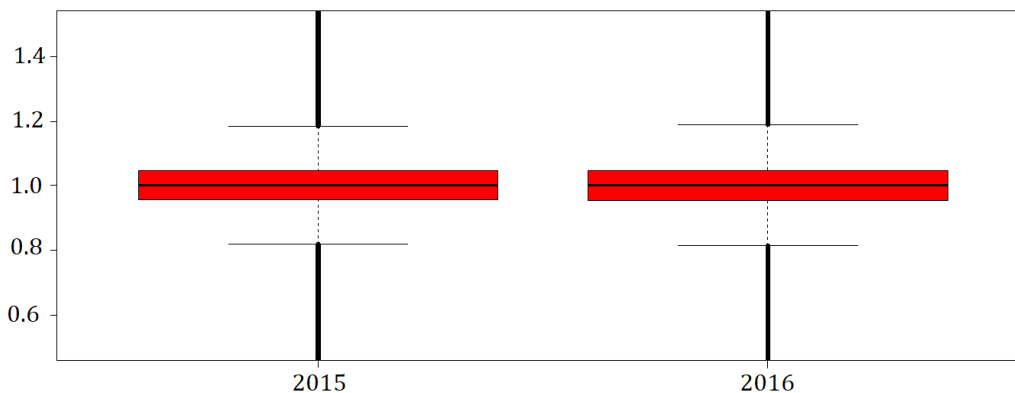
Boxplot anual de coeficiente de sub-declaración



Fuente: Elaboración propia con información del SAT.

Al limitar el eje vertical con valores entre 0.5 y 1.5, se puede visualizar el siguiente comportamiento.

Boxplot ajustado de coeficiente de sub-declaración



Fuente: Elaboración propia con información del SAT.

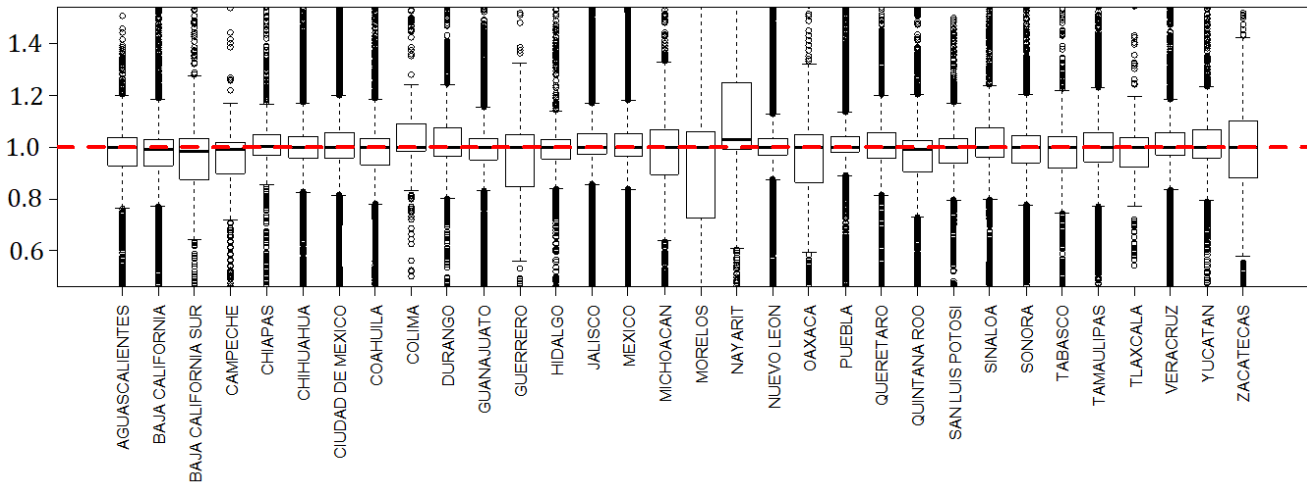
También se puede asegurar que hay un 50% que tienen un cociente por encima de la unidad y otro 50% por debajo de este valor. Además, se observa simetría en las cajas, lo cual nos indica que existe un comportamiento similar de los contribuyentes que tienen indicios de sub-declaración o sub-facturación. Sólo en los casos extremos (valores altos y bajos) es donde las diferencias son amplias.

Al realizar la gráfica boxplot del cociente de sub-declaración por estado, se obtiene un comportamiento por entidad alrededor de uno, la amplitud de la caja en cada estado nos indica que existe mayor o menor variabilidad (el hecho de que exista variabilidad en el cociente es que los contribuyentes tienden a sub-declarar o sub-facturar con mayor frecuencia).

Caso Personas Morales

Los contribuyentes registrados en los estados que presentan el cuadro superior de la gráfica boxplot más amplia, indican que los contribuyentes registrados en esa entidad tienden a tener un índice de sub-declaración mayor al de sub-facturación. Por ejemplo, los contribuyentes registrados en el estado de Nayarit presentan mayor variabilidad en su coeficiente de sub-declaración.

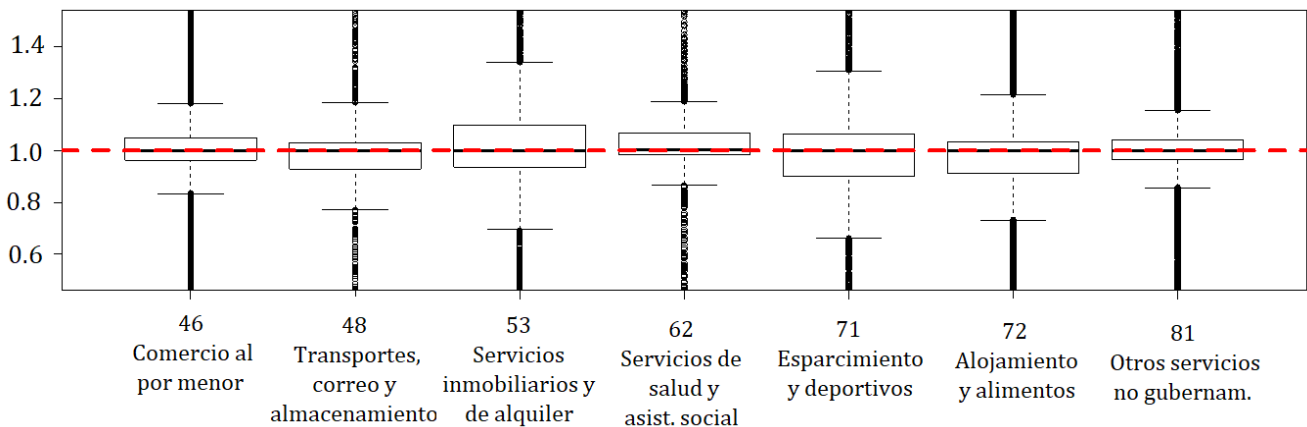
Boxplot ajustado de coeficiente de sub-declaración por entidad federativa



Fuente: Elaboración propia con información del SAT.

Por sector se tiene una mayor variabilidad en el sector 53 de servicios inmobiliarios y de alquiler.

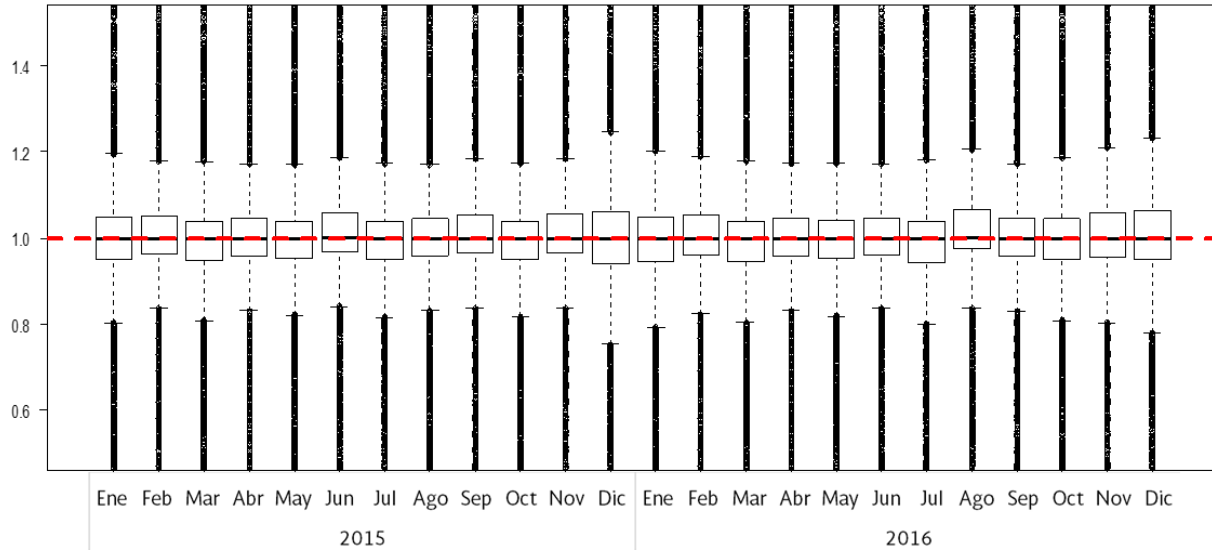
Boxplot ajustado de coeficiente de sub-declaración por actividad económica



Fuente: Elaboración propia con información del SAT.

Desde la perspectiva temporal se realizó el mismo análisis por mes, lo que muestra que en el mes de diciembre los contribuyentes tienden a sub-declarar al final del año, dado que la caja es más amplia por la parte superior.

Boxplot ajustado de coeficientes por mes

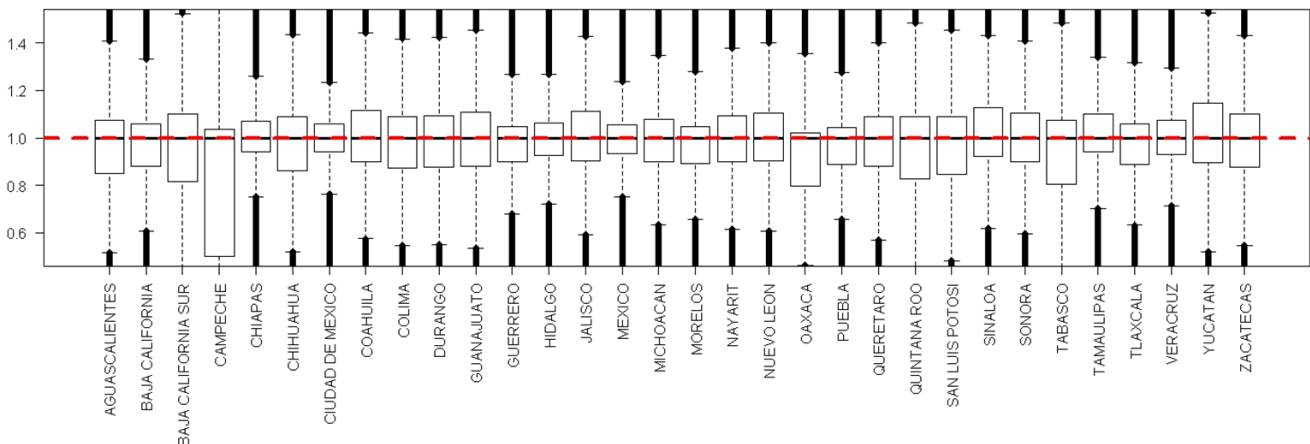


Fuente: Elaboración propia con información del SAT.

Caso Personas Físicas

Los contribuyentes que están registrados en los estados de Coahuila, Sinaloa y Yucatán, presentan una caja superior a la mediana con mayor altura, es decir, estos contribuyentes tienden a tener un índice de sub-declaración mayor. En sentido contrario, los contribuyentes registrados en Campeche, Oaxaca y Tabasco, presentan mayor sub-facturación.

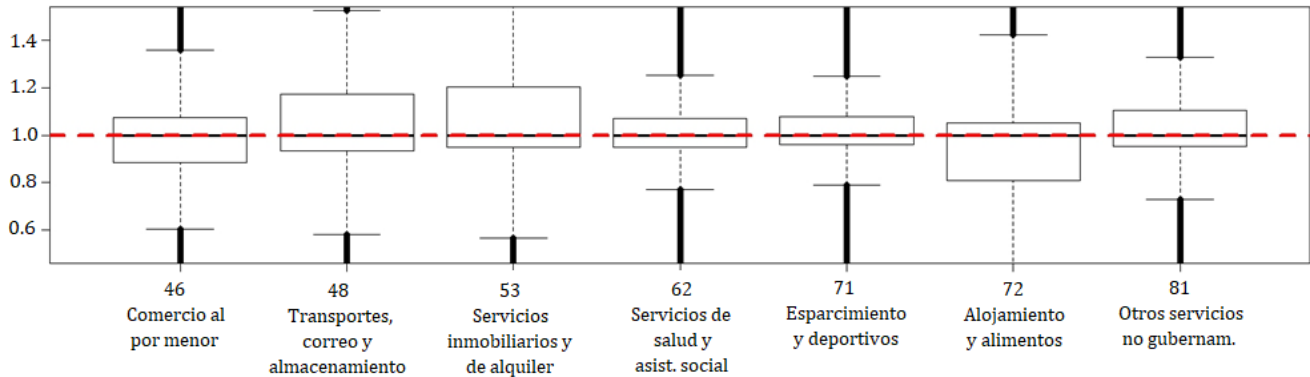
Boxplot ajustado de coeficiente de sub-declaración por entidad federativa



Fuente: Elaboración propia con información del SAT.

Por sector se tiene una mayor variabilidad en la sub-declaración en el sector 53 de “Servicios inmobiliarios y de alquiler”, mientras que, en el sector 72 de servicios de “Alojamiento y alimentos”, presenta una mayor variabilidad con respecto a la sub-facturación.

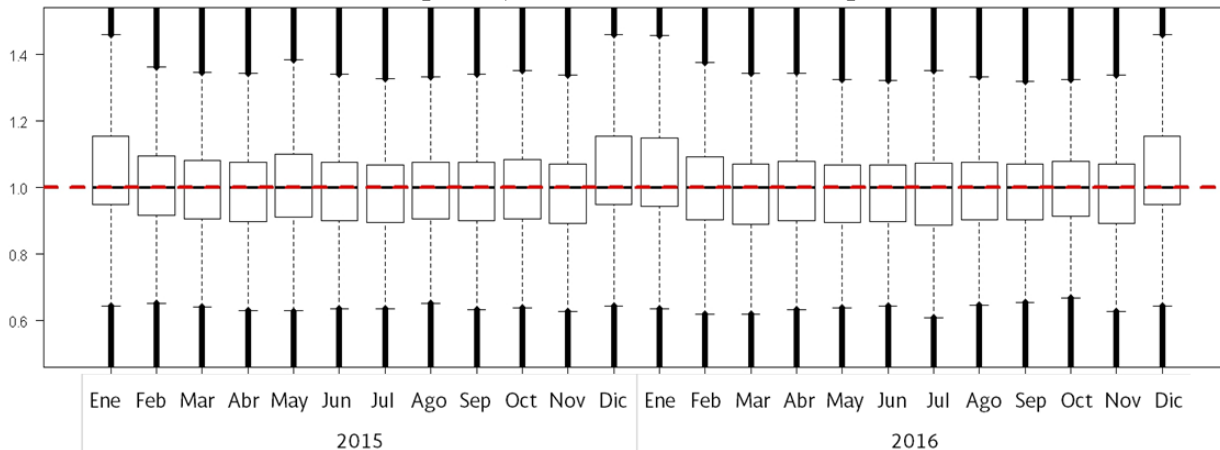
Boxplot ajustado de coeficiente de sub-declaración por actividad económica



Fuente: Elaboración propia con información del SAT.

Para el análisis mensual, se presenta el mismo comportamiento de sub-declaración en el mes de diciembre, pues presenta las cajas superiores más altas.

Boxplot ajustado de coeficientes por mes

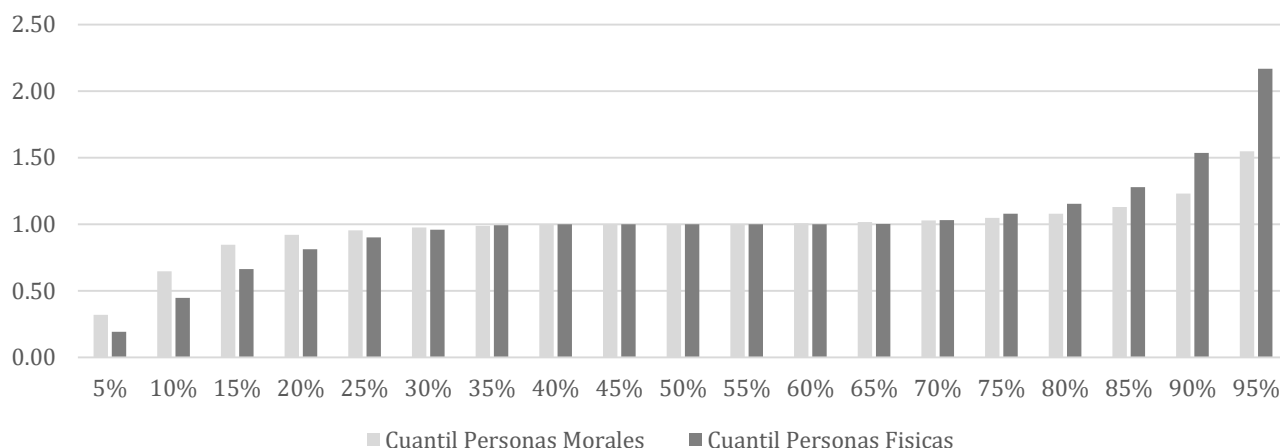


Fuente: Elaboración propia con información del SAT.

Detección de Evasión

Para determinar que puntos de corte son relevantes y tener indicios de que existe sub-declaración o sub-facturación, se realiza la siguiente tabla de cuantiles para el cociente de sub-declaración, calculado para personas morales y físicas. Con base en lo anterior, se observa que los cuantiles del 20% y del 80% son lo que presentan un valor del cociente significativamente diferente a uno.

Cuantiles del cociente de Sub-declaración



Fuente: Elaboración propia con información del SAT.

Por lo tanto, las personas morales que presentan un cociente de sub-declaración mayor a 1.0792, son significativamente diferentes a uno y presentan sub-declaración de ingresos. Mientras que, si es menor a 0.9206 es significativamente diferente de uno y presenta sub-facturación.

En el caso de personas físicas, se está sub-declarando si se presenta un cociente mayor a 1.1531 y sub-facturando si es menor a 0.8123, estos márgenes de holgura se estiman debido a que pueden existir diversos factores que afecte la exacta igualdad a la unidad.

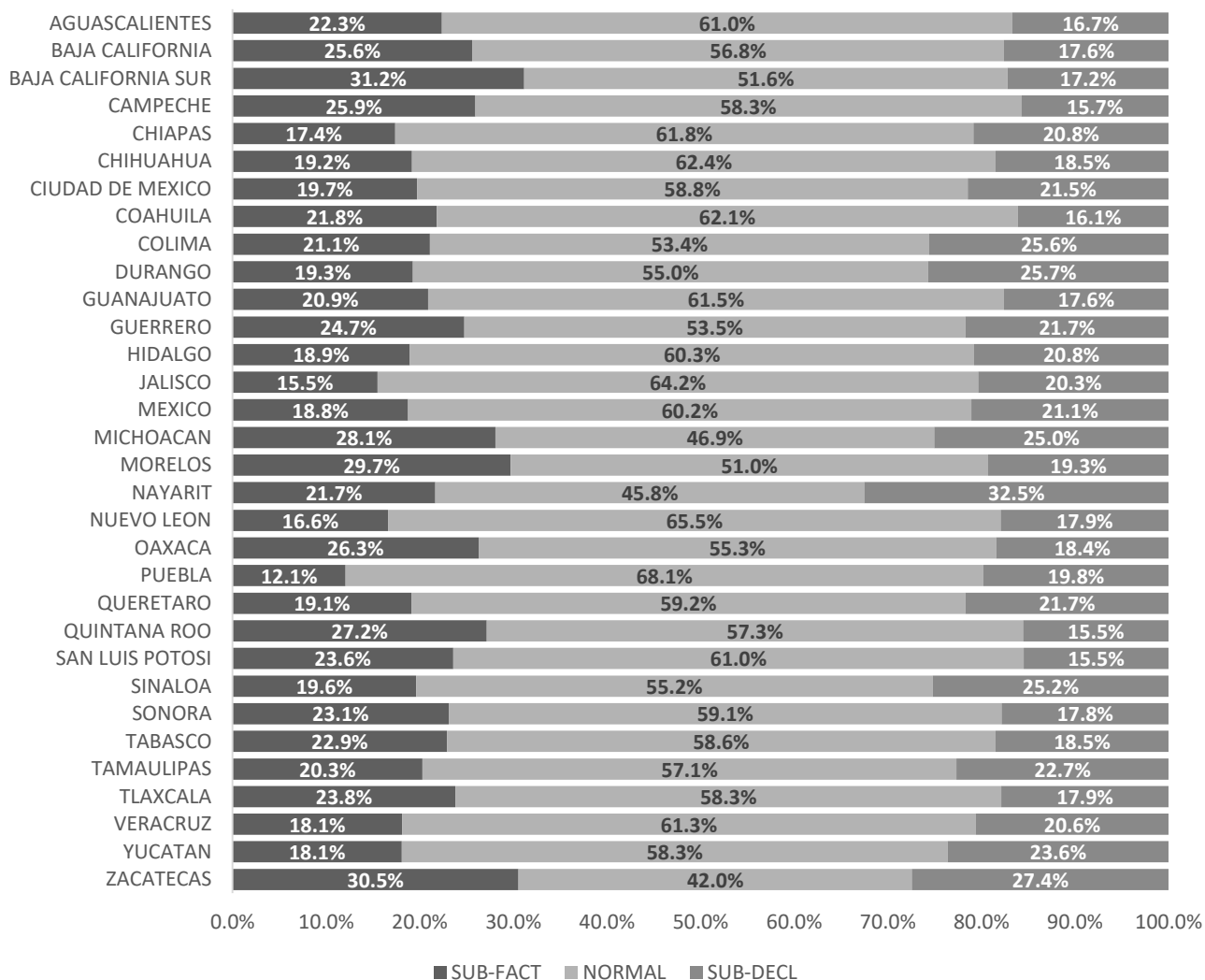
Los contribuyentes que se encuentran sub-declarando ingresos, dado el criterio anterior corresponde al 20.16%⁴. La entidad Federativa con mayor incidencia de contribuyentes que sub-declaran se encuentran registrados en Nayarit con el 27.5% de la muestra analizada en esa entidad. Mientras que los registrados en Campeche presenta la menor incidencia con 14.2%.

⁴ Este hecho coincide con el principio de Pareto en la que aproximadamente el 20% de la población genera, en este caso, el 80% de las desviaciones.

Análisis Personas Morales

Los contribuyentes que se encuentran sub-declarando ingresos, dado el criterio anterior corresponde al 20.17%. La entidad Federativa con mayor incidencia de contribuyentes que sub-declaran se encuentran registrados en Nayarit con el 32.5% de la muestra analizada en esa entidad, mientras que los registrados en San Luis Potosí presenta la menor incidencia con 15.5%.

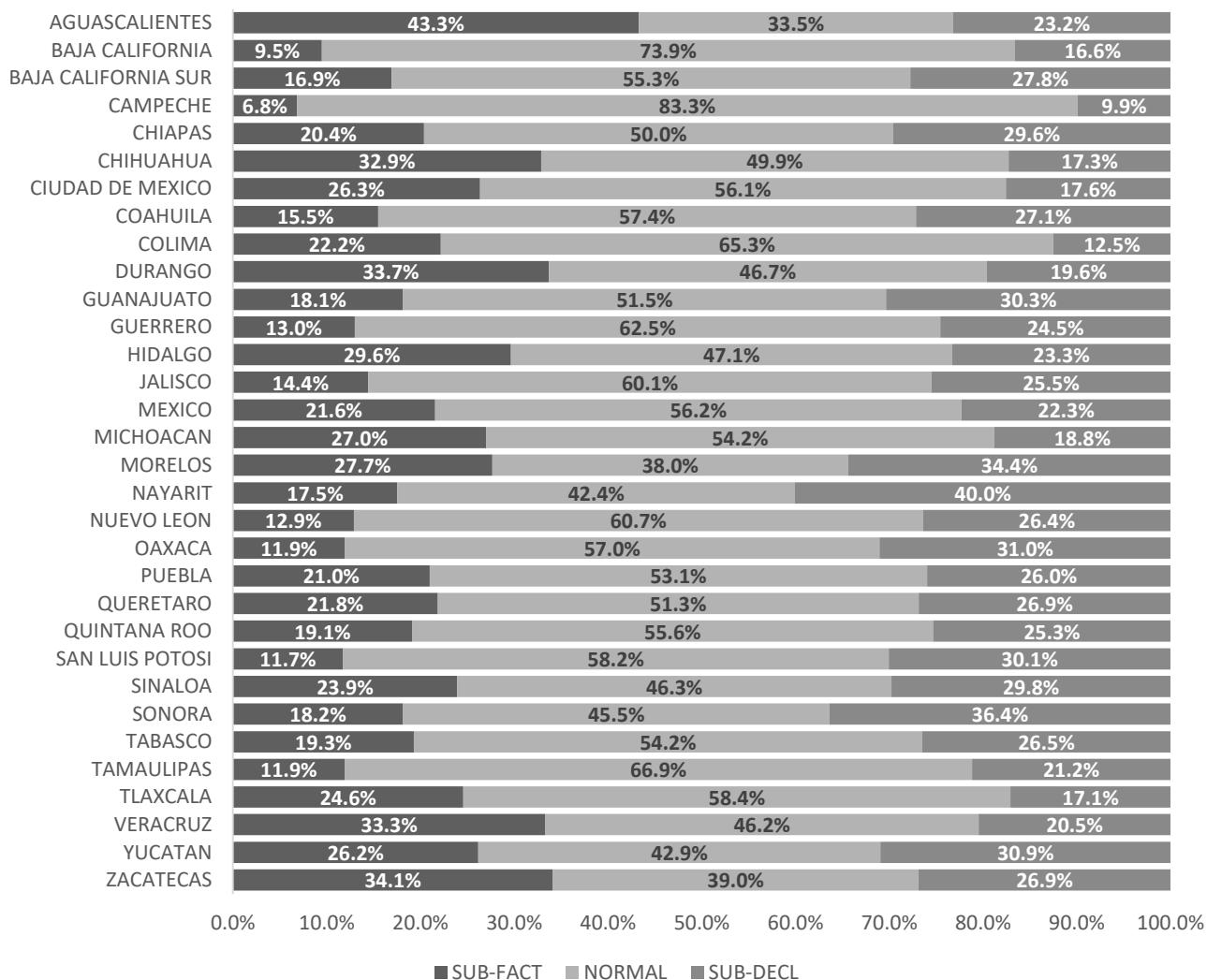
Participación de contribuyentes por Entidad Federativa



Fuente: Elaboración propia con información del SAT.

En términos del porcentaje de ingreso potencial para cada estado, igualmente Nayarit presenta una concentración mayor en ingresos no declarados por los contribuyentes registrados en esa entidad.

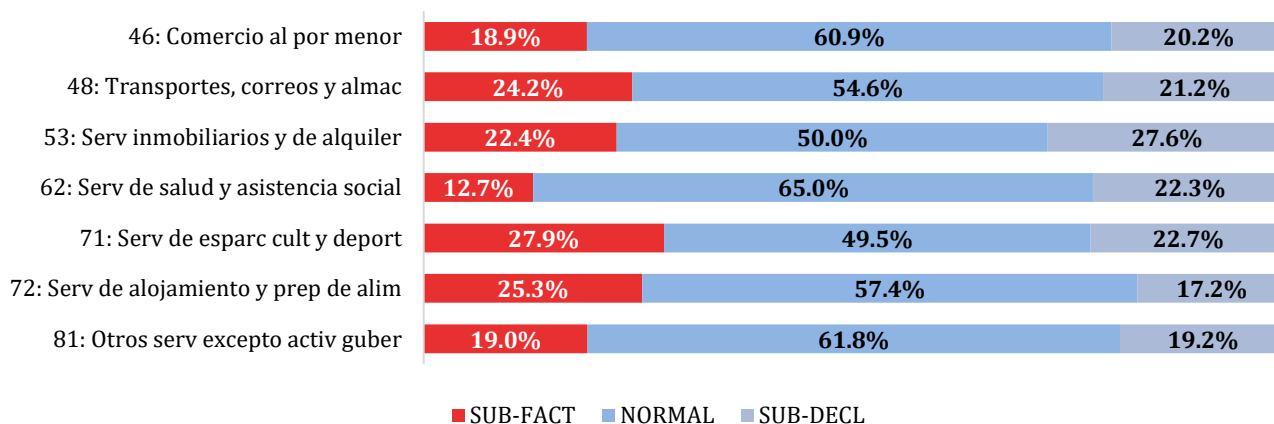
Participación del ingreso potencial por Entidad Federativa



Fuente: Elaboración propia con información del SAT.

Dentro de los sectores de actividad económica, los contribuyentes que muestran una mayor incidencia por sub-declaración, se encuentran en el sector de servicios inmobiliarios y de alquiler con un 27.6% de contribuyentes que superan el cociente de 1.08, mientras que los de menor incidencia son las empresas de otros servicios de alojamiento y preparación de alimentos.

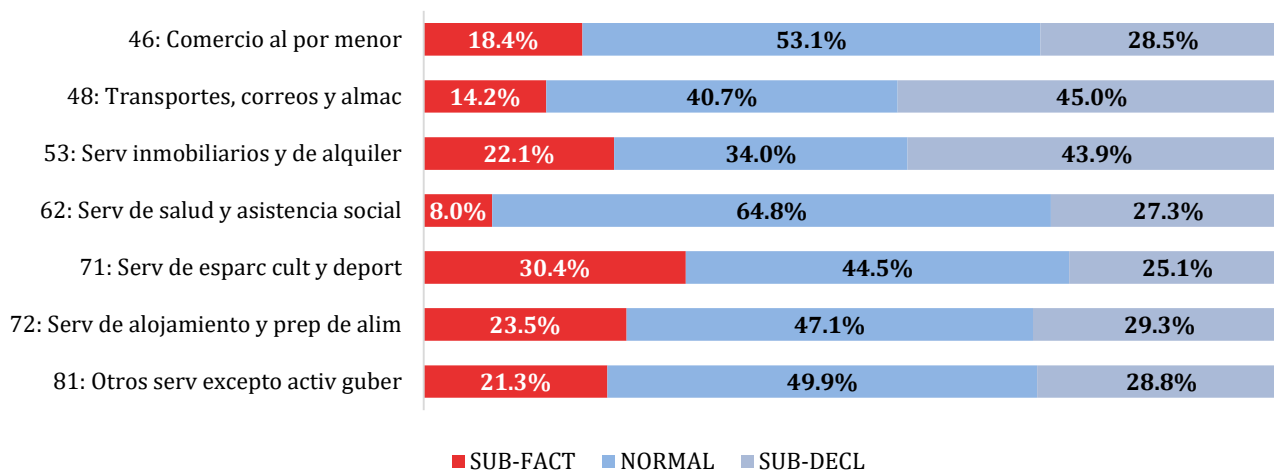
Participación de contribuyentes por actividad económica



Fuente: Elaboración propia con información del SAT.

En términos del porcentaje de ingreso potencial, el sector de transporte, correos y almacenamiento, presenta mayores montos sub-declarados (45%).

Participación del ingreso potencial por actividad económica

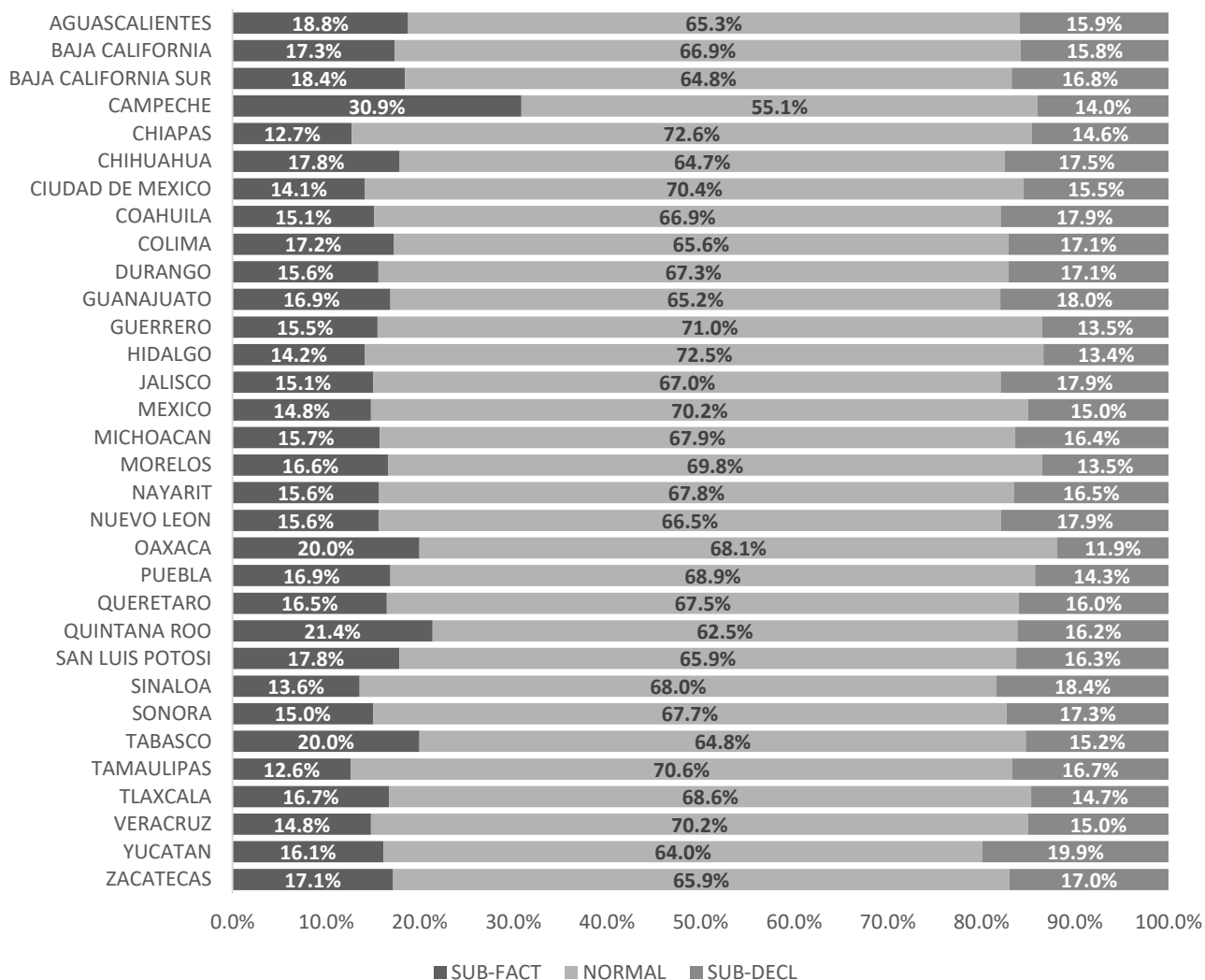


Fuente: Elaboración propia con información del SAT.

Análisis Personas Físicas

Los contribuyentes que se encuentran sub-declarando ingresos, dado el criterio anterior corresponde al 16.12%. La entidad Federativa con mayor incidencia de personas físicas que sub-declaran se encuentran registradas en Yucatán con el 19.9% de la muestra analizada. Mientras que los registrados en Oaxaca presenta la menor incidencia con 11.9%.

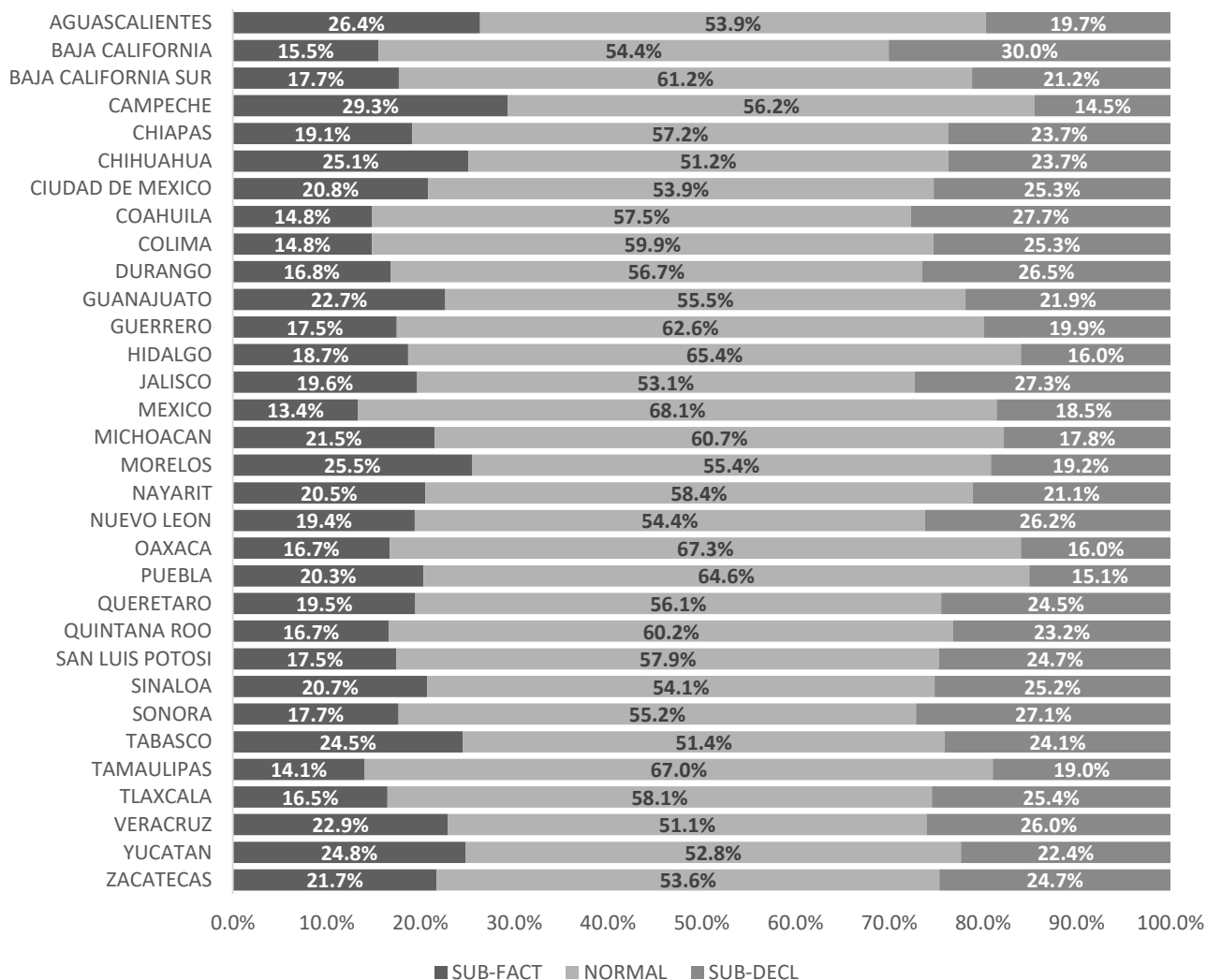
Participación de contribuyentes por Entidad Federativa



Fuente: Elaboración propia con información del SAT.

En términos del porcentaje de ingreso potencial para cada estado, los contribuyentes registrados en Baja California presentan el 30% de los ingresos no declarados, mientras que los registrados en Campeche representan el 14.5%.

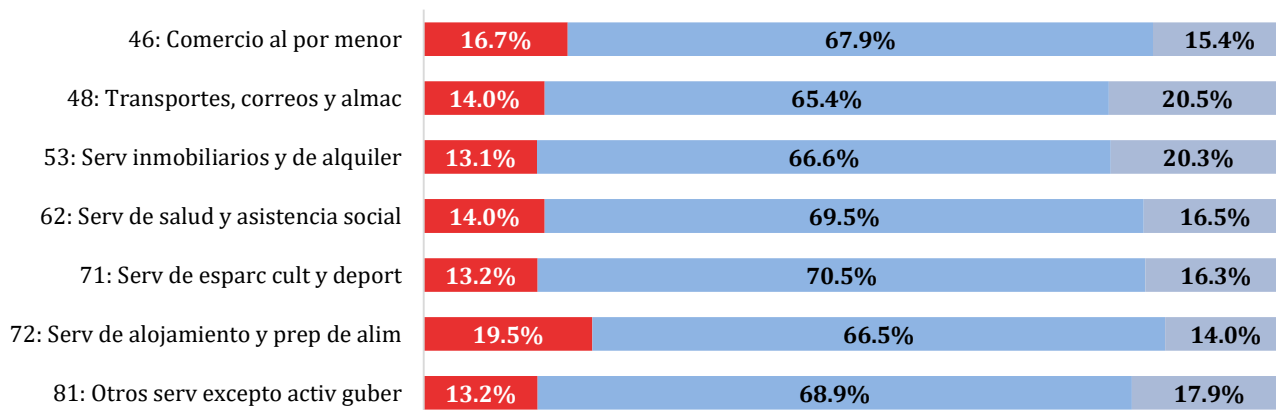
Participación del ingreso potencial por Entidad Federativa



Fuente: Elaboración propia con información del SAT.

Con respecto a los sectores de actividad económica las empresas que tienen indicios de tener sub-declaración en sus ingresos, presentan mayor densidad en el sector de Transportes, correos y almacenamiento con 20.5% de contribuyentes que superan el cociente de 1.15, mientras que los de menor incidencia son las empresas de otros servicios de alojamiento y preparación de alimentos con 14%.

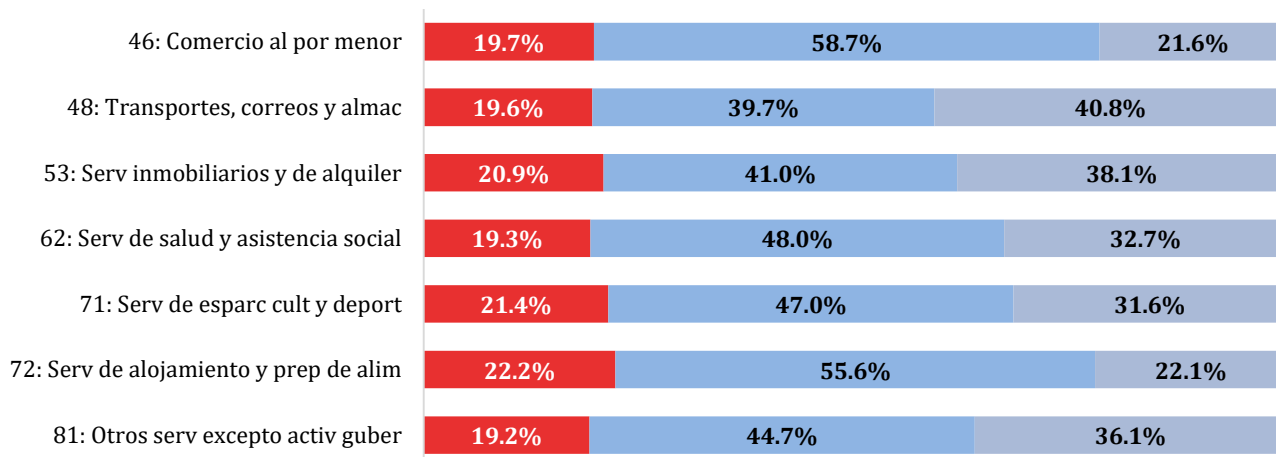
Participación de contribuyentes por actividad económica



Fuente: Elaboración propia con información del SAT.

En términos del porcentaje de ingreso potencial para cada sector, también el sector de transportes, correos y almacenamiento presenta la mayor incidencia en sub-declaración de ingresos con 40.8%.

Participación del ingreso potencial por actividad económica



Fuente: Elaboración propia con información del SAT.

4. Análisis de Evasión

Para el cálculo de evasión se asumirá que, los contribuyentes que presentan sub-declaración por encima del cociente 1.0792 para personas morales y 1.15 para personas físicas son los que están evadiendo. El impuesto de ISR no declarado, será calculado a partir del ingreso potencial obtenido, aplicando las reglas contables correspondientes. En este sentido, la evasión del ISR se obtiene con base al impuesto causado, debido a que obtener el impuesto recaudado implicaría un mayor número de supuestos.

$$Tasa\ de\ Evasi\o{n} = 100 * \left[\frac{Impuesto\ causado\ potencial - Impuesto\ causado\ declarado}{Impuesto\ causado\ potencial} \right]$$

De las 4,568 personas morales que se analizaron, se observa que el porcentaje de ingreso no declarado se ha incrementado en 0.6%, y que la evasión de ISR ha disminuido en 0.2 puntos del 2015 a 2016. Sí sólo se consideran a los contribuyentes que sub-declaran ingresos, la evasión ha disminuido en 3.1 puntos porcentuales en el mismo periodo.

Resumen de resultados, personas morales (millones de pesos)

Concepto	2015	2016
Ingreso Potencial	860,577.7	961,231.8
Declaraciones y Pagos	768,435.1	853,949.7
% de Ingreso no declarado	12.0%	12.6%
Total de contribuyentes		
ISR Potencial	10,423.9	12,474.1
Evasi\o{n} de ISR	1,084.8	1,273.7
% de Evasi\o{n} de ISR	10.4%	10.2%
Empresas que sub-declaran		
ISR Potencial	2,923.7	3,751.8
Evasi\o{n} de ISR	1,084.8	1,273.7
% de Evasi\o{n} de ISR	37.1%	34.0%

Fuente: Elaboraci\o{n} propia con informaci\o{n} del SAT.

De las cerca de 100 mil personas f\xedsicas de la muestra analizada, se observa que el porcentaje de ingreso no declarado se ha mantenido pr\acticamente constante alrededor del 9.8%, el porcentaje de evasi\o{n} del ISR ha disminuido en 9.7 puntos porcentuales. Considerando \uunicamente las empresas que sub-declaran, la evasi\o{n} ha disminuido 3.8 puntos porcentuales de 2015 a 2016.

Resumen de resultados, personas físicas
(millones de pesos)

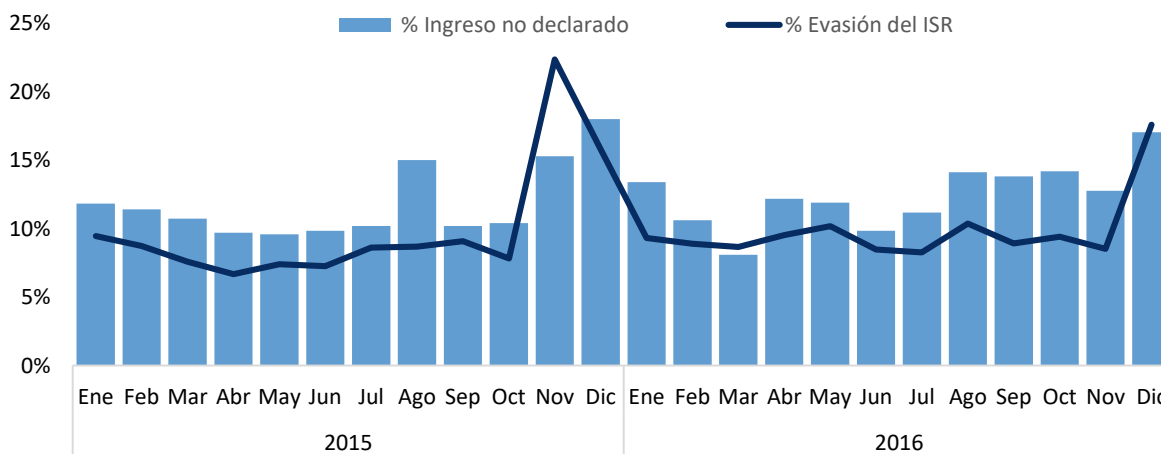
Concepto	2015	2016
Ingreso Potencial	299,983.7	316,613.3
Declaraciones y Pagos	272,984.9	288,351.5
% de Ingreso no declarado	9.9%	9.8%
Total de contribuyentes		
ISR Potencial	12,017.1	14,742.5
Evasión de ISR	8,099.2	8,511.0
% de Evasión de ISR	67.4%	57.7%
Empresas que sub-declaran		
ISR Potencial	8,675.7	9,505.6
Evasión de ISR	8,099.2	8,511.0
% de Evasión de ISR	93.4%	89.5%

Fuente: Elaboración propia con información del SAT.

Caso Personas Morales

El comportamiento de los ingresos no declarados para 2015 muestra repuntes en los meses de agosto, noviembre y diciembre. En el caso de noviembre y diciembre, los repuntes en los ingresos no declarados se reflejan en los montos de ISR causado que se evade (22.4% y 15.8% respectivamente). Sin embargo, en agosto dicho repunte no se refleja en la evasión del ISR, lo que pareciera una inconsistencia en la información.

Ingresos no declarados y evasión mensual del ISR, personas morales



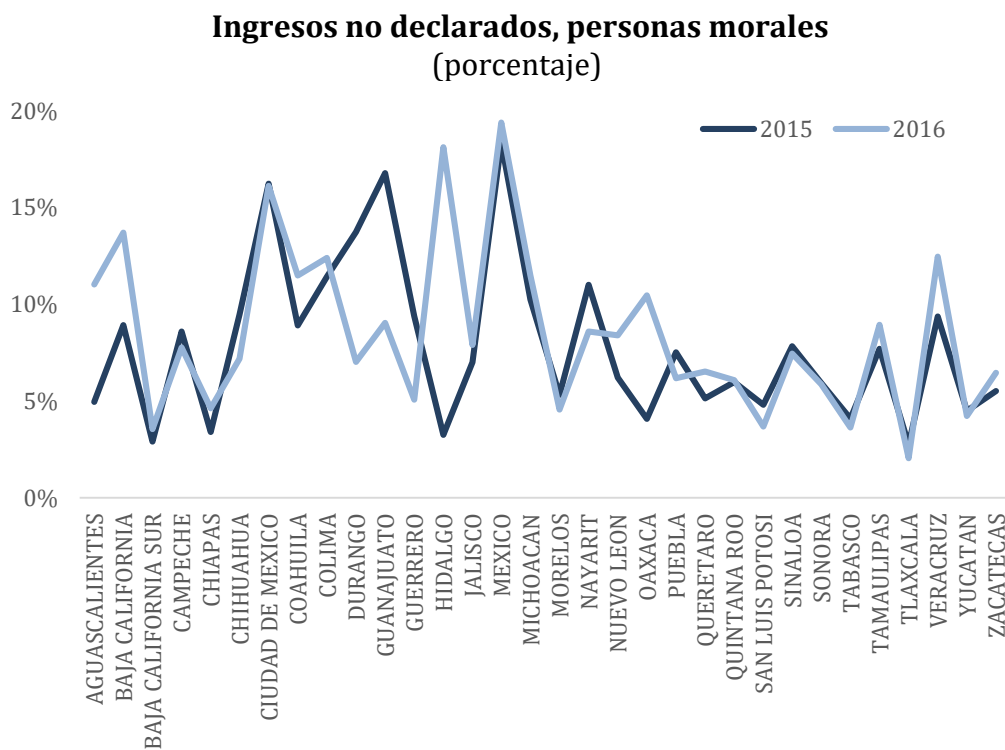
Fuente: Elaboración propia con información del SAT.

En 2016, los meses de agosto y diciembre vuelven a mostrar repuntes en el porcentaje de ingresos no declarados, en este año, de igual manera se presenta un alza en la tasa de evasión del ISR para estos meses.

Resultados por Entidad Federativa

Como se esperaba, el ingreso potencial por estados de la república está concentrado en primer lugar, en la Ciudad de México, dado que muchos de los corporativos de empresas con cobertura nacional se establecen allí, después le siguen Jalisco, Nuevo León y Estado de México en nivel de importancia.

En cuanto a los ingresos no declarados, para 2015 se tiene una media de 12% y para 2016 de 12.6%. Para 2016, los contribuyentes registrados en Estado de México, Hidalgo y la Ciudad de México, son los que tiene un mayor porcentaje de ingresos no declarados.



Fuente: Elaboración propia con información del SAT.

En 2015, los contribuyentes registrados en Nayarit y Ciudad de México presentan los mayores niveles de evasión con 22% y 17.8% respectivamente. Por otro lado, los contribuyentes registrados en los estados de Zacatecas y Tlaxcala tienen las tasas de evasión más bajas con 0.9% y 2.9% respectivamente.

**Tasa de evasión por entidad federativa, personas morales
(porcentaje)**

Entidades	2015	2016	Entidades	2015	2016
Coahuila	12.0	32.1	Querétaro	3.7	6.5
Nayarit	22.0	15.9	Tabasco	5.6	6.4
Veracruz	8.5	15.9	Sonora	6.4	6.3
Aguascalientes	4.7	12.7	Baja California Sur	3.3	6.2
Hidalgo	4.1	12.3	San Luis Potosí	5.2	6.2
Ciudad de México	17.8	12.3	Morelos	5.0	6.1
Baja California	10.3	12.2	Puebla	8.2	5.7
Estado de México	10.7	11.6	Sinaloa	5.8	5.6
Campeche	5.4	10.9	Yucatán	5.8	5.3
Tamaulipas	5.7	9.3	Quintana Roo	9.6	5.2
Guanajuato	8.3	8.8	Chiapas	3.3	5.1
Michoacán	13.5	8.3	Chihuahua	8.0	5.0
Colima	9.2	8.2	Durango	7.2	5.0
Nuevo León	6.1	8.1	Guerrero	14.0	4.0
Jalisco	6.6	8.1	Tlaxcala	2.9	3.8
Oaxaca	3.8	7.9	Zacatecas	0.9	2.0
			Total	10.4	10.2

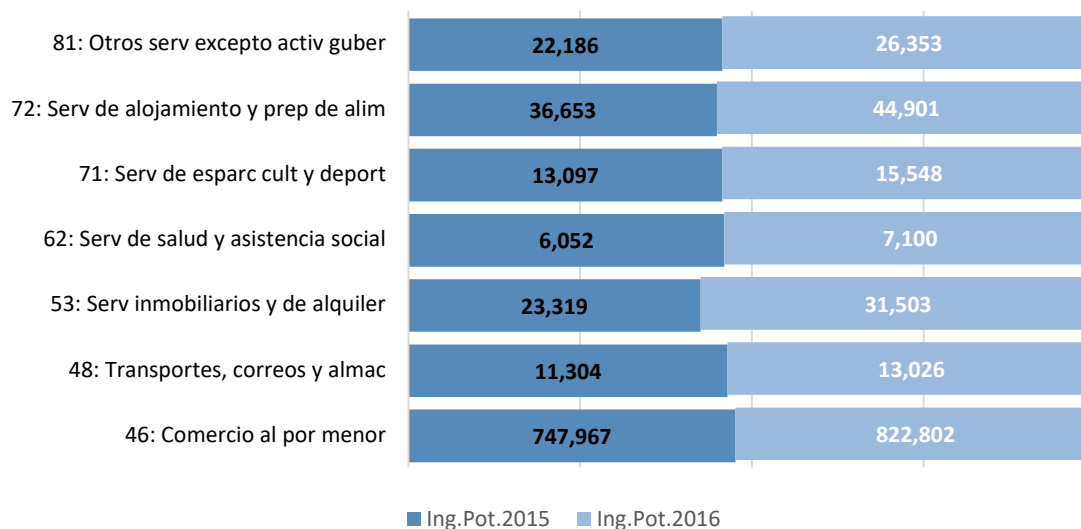
Fuente: Elaboración propia con información del SAT.

Para 2016, los contribuyentes registrados en las entidades de Coahuila y Nayarit, muestran las tasas de evasión más altas con 32.1% y 15.9%, respectivamente. Mientras que los contribuyentes de Zacatecas y Tlaxcala mantienen las menores tasas de evasión.

Resultados por Sector Económico

De los sectores económicos seleccionados, se observa que el “Comercio al por menor” concentra la mayor parte del ingreso potencial, seguido por el sector de “Servicios de alojamiento y preparación de alimentos”.

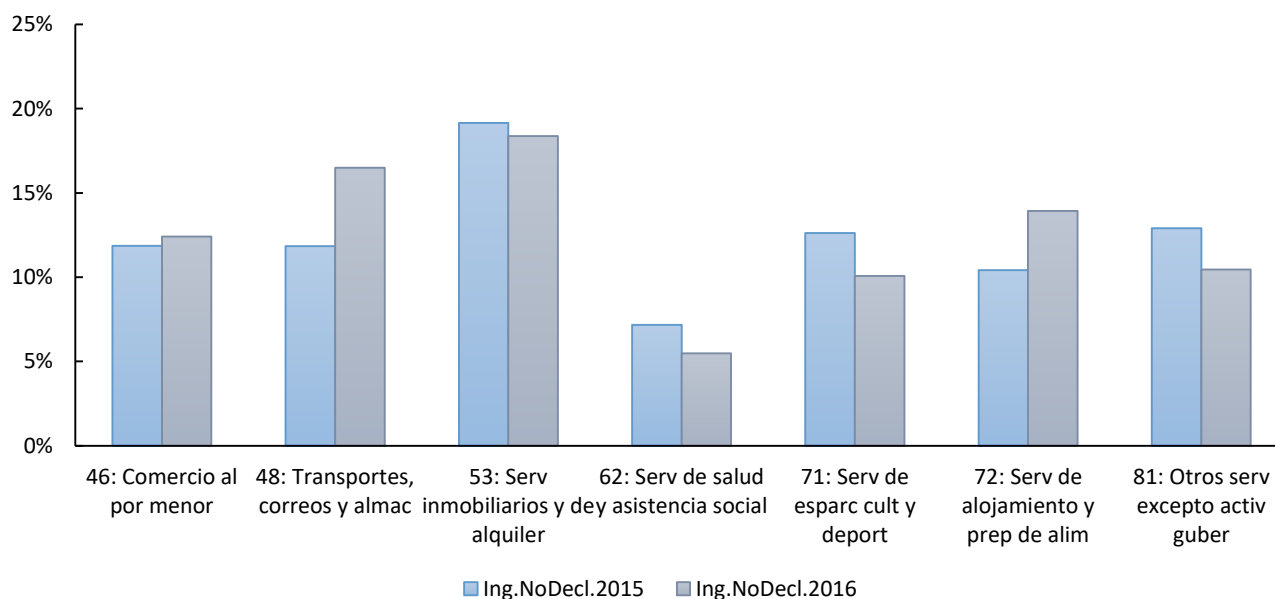
Ingreso potencial por actividad económica, personas morales (millones de pesos)



Fuente: Elaboración propia con información del SAT.

El porcentaje de ingreso no declarado más alto se tiene en el sector de Servicios inmobiliarios y de alquiler consistentemente ambos años, el sector de Transportes, correos y almacenamiento tuvo un incremento importante entre el 2015 y 2016, el resto de sectores ha permanecido estable o con una reducción en los ingresos no declarados para 2016 respecto a 2015.

Ingresos no declarados por sector de actividad económica, personas morales (porcentaje)



Fuente: Elaboración propia con información del SAT.

El sector de “Servicios inmobiliarios y de alquiler”, tanto para 2015 como para 2016, presenta la evasión más alta con respecto a los demás sectores, seguida del sector “Transportes, correo y almacenamiento”.

Tasas de evasión por sector de actividad económica, personas morales (porcentaje)

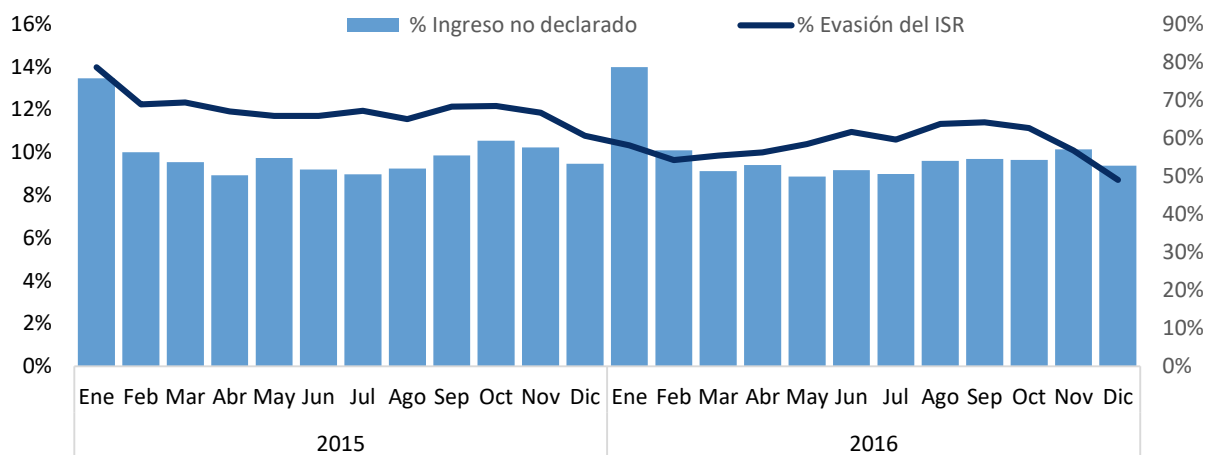
Sector de actividad económica	2015	2016
53: Servicios inmobiliarios y de alquiler	28.1	24.6
48: Transportes, correo y almacenamiento	16.6	15.7
72: Alojamiento y alimentos	7.1	12.3
81: Otros servicios no gubernamentales	13.8	10.4
46: Comercio al por menor	8.2	8.9
71: Esparcimiento y deportivos	9.9	6.3
62: Servicios de salud y asistencia social	5.9	5.5
TOTAL	10.4	10.2

Fuente: Elaboración propia con información del SAT.

Caso Personas Físicas

Las personas físicas, muestran repuntes en sus ingresos no declarados en el mes de enero, a diferencia de las personas morales que lo presentaban al final del año. En cuanto al comportamiento de la evasión, se observa una tendencia decreciente de enero 2015 a diciembre 2016.

Ingresos no declarados y evasión mensual del ISR, personas físicas



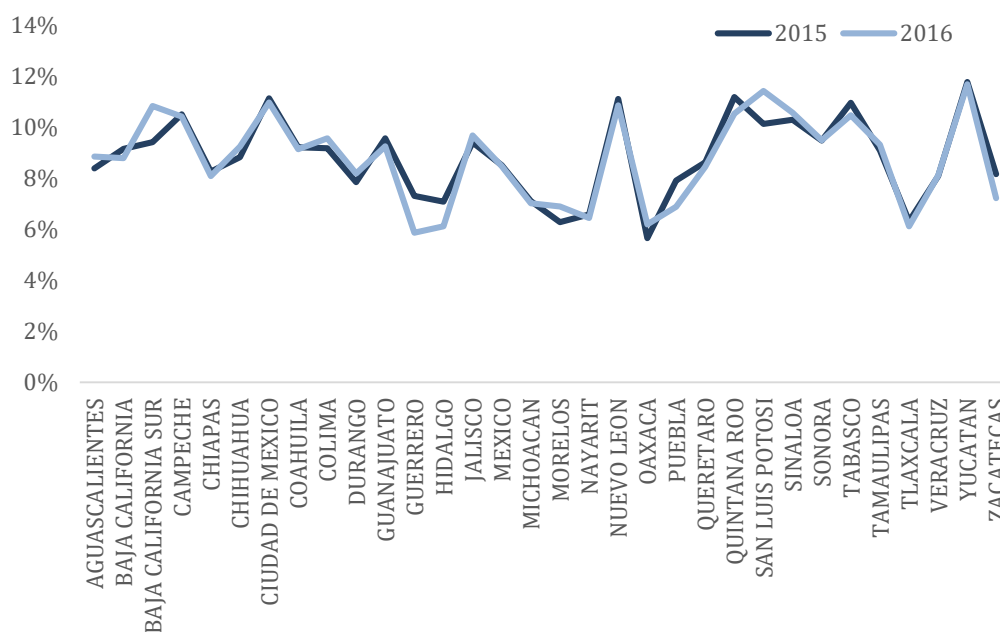
Fuente: Elaboración propia con información del SAT.

Resultados por Entidad Federativa

Como se esperaba, el ingreso potencial de los contribuyentes registrados en los estados de la república está concentrado en primer lugar, en la Ciudad de México, dado que la mayoría de personas físicas con actividad empresarial están registradas en esta entidad.

En lo que respecta a los ingresos no declarados, tanto en 2015 como en 2016, los estados se comportan de manera estable, manteniendo las mismas tasas. Las personas físicas registradas en Yucatán presentan ingresos no declarados más altos (11.7% con respecto al potencial), seguidos por los contribuyentes de los estados de San Luis Potosí (11.4%), Ciudad de México (11.0%), Nuevo León (10.9%) y Baja California (10.8%).

Ingresos no declarados, personas físicas



Fuente: Elaboración propia con información del SAT.

En general los contribuyentes analizados por cada Entidad Federativa, presentan una disminución en su tasa de evasión. En total la evasión disminuyó 9.7 puntos porcentuales, al pasar de 67.4% en 2015 a 57.7% para 2016, la entidad que reportó el mayor decremento en su tasa de evasión fue Guerrero, con 16.7%. Los contribuyentes registrados en el estado de Yucatán, presentaron una disminución de 5% en su tasa de evasión, sin embargo, continúa siendo los que presentan mayores tasas de evasión.

**Tasa de evasión por entidad federativa, personas físicas
(porcentaje)**

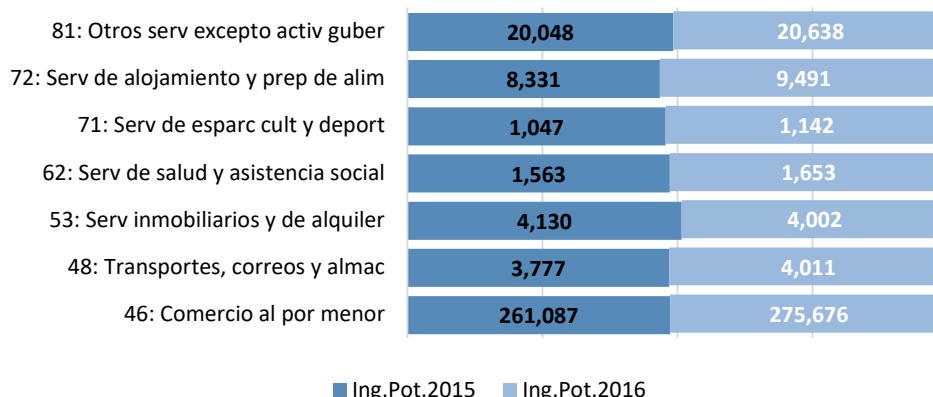
Entidades	2015	2016	Entidades	2015	2016
Yucatán	74.1	69.1	Tamaulipas	65.4	56.8
Colima	68.0	65.2	Puebla	70.4	56.8
Nuevo León	71.6	62.9	Durango	69.6	56.6
Quintana Roo	68.8	62.7	Baja California	64.3	56.5
Guanajuato	73.1	62.3	Campeche	66.5	56.1
San Luis Potosí	71.2	61.8	Tlaxcala	59.6	56.0
Sinaloa	69.1	61.2	Chiapas	66.8	56.0
Chihuahua	69.6	60.3	Zacatecas	67.2	54.5
Ciudad de México	66.6	59.6	Estado de México	66.0	54.0
Jalisco	70.5	59.2	Michoacán	65.8	53.5
Tabasco	69.8	58.7	Morelos	60.1	53.4
Veracruz	64.5	58.3	Querétaro	62.8	52.6
Aguascalientes	66.7	57.4	Nayarit	65.0	51.6
Coahuila	67.2	57.4	Oaxaca	54.2	50.9
Sonora	66.9	57.0	Hidalgo	63.7	49.8
Baja California Sur	62.1	56.9	Guerrero	64.7	48.0
			Total	67.4	57.7

Fuente: Elaboración propia con información del SAT.

Resultados por Sector Económico

El ingreso potencial por sector está fuertemente concentrado en Comercio al por menor seguido por Servicios de alojamiento y preparación de alimentos. Estos resultados se presentan tanto en personas morales como personas físicas, derivado de los criterios considerados para la selección de la muestra.

Ingreso potencial por actividad económica, personas físicas (millones de pesos)

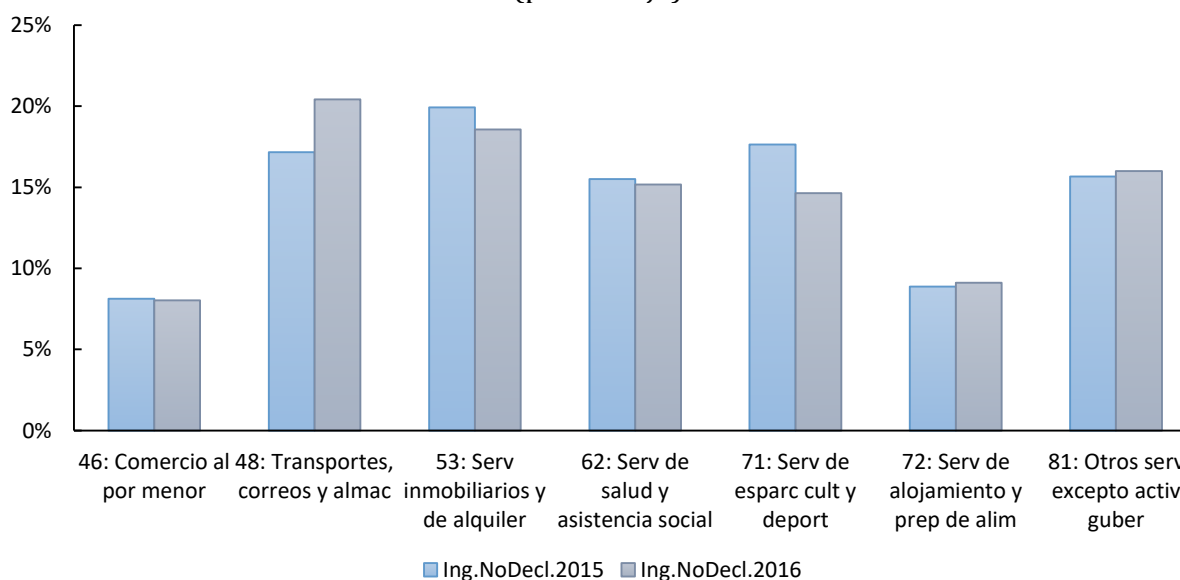


Fuente: Elaboración propia con información del SAT.

El porcentaje de ingreso no declarado más alto se observa en el sector de “Transportes, correos y almacenamiento” (20.4% para 2016), el cual tuvo un incremento importante entre el 2015 y 2016 (3.3%). El sector “Servicios inmobiliarios y de alquiler” presenta consistentemente en ambos años porcentajes de ingreso no declarado similares (19.9% para 2015 y 18.6% en 2016).

El sector de “Servicios de esparcimiento cultural y deportivos”, muestra una disminución en el porcentaje de ingresos no declarados, al pasar de 17.6% en 2015 a 14.6% para 2016. El resto de sectores ha permanecido estable o con variaciones no significativas.

Ingresos no declarados por sector de actividad económica, personas físicas (porcentaje)



Fuente: Elaboración propia con información del SAT.

El sector de “Transporte, correos y almacenamiento”, muestra las tasas más altas de evasión, seguido de “Otros servicios no gubernamentales”. Todos los sectores presentan decremento en la tasa de evasión del 2015 al 2016.

Tasas de evasión por sector de actividad económica, personas físicas
(porcentaje)

Sector de actividad económica	2015	2016
48: Transportes, correo y almacenamiento	72.8	69.7
81: Otros servicios no gubernamentales	70.3	64.9
53: Servicios inmobiliarios y de alquiler	70.3	61.7
46: Comercio al por menor	67.5	56.9
62: Servicios de salud y asistencia social	54.4	51.6
72: Alojamiento y alimentos	55.5	49.5
71: Esparcimiento y deportivos	61.1	47.9
TOTAL	67.4	57.7

Fuente: Elaboración propia con información del SAT.

5. Conclusiones y recomendaciones

La tendencia mundial en la sustitución de documentos impresos por medios electrónicos ha revolucionado los procesos de las administraciones tributarias. Dos frentes que se han desarrollado casi paralelamente son: por un lado, la recepción electrónica de las diversas declaraciones y por otro, el registro en tiempo real de cada operación comercial mediante el uso de la factura electrónica.

Estas dos fuentes de información constituyeron la base del presente estudio, el cual calculó el potencial de recaudación de los contribuyentes que en su actividad económica suelen emitir tickets. Para dicho cálculo, se procesó la información contenida en cerca de mil millones de facturas electrónicas emitidas por personas morales del régimen general y personas físicas con actividad empresarial correspondiente a los ejercicios 2015 y 2016.

Se utilizó el análisis descriptivo y de aprendizaje automático para determinar las características de los contribuyentes que tienen indicios a sub-declarar y a sub-facturar. Se establece un cociente de sub-declaración contrastando los ingresos mensuales por facturas CFDI y el ingreso nominal realizado en declaración mensual de ingresos y pagos.

La información proveniente de CFDI fue sometida a un proceso de reclasificación y limpieza en los campos de “Método de pago” y “Tipo de Moneda” ya que en su mayoría los contribuyentes utilizan diversas formas para estos campos. Así mismo, se utilizaron técnicas de búsquedas de texto para clasificarlas en pesos o monedas de diversos países.

Para la clasificación de los contribuyentes que sub-declaran, sub-facturan o declaran normalmente, se realizaron pruebas de hipótesis estadística para igualdad de medias en cada contribuyente, respecto a sus declaraciones. Posteriormente se construyó un árbol de decisión que permitió clasificar a cada uno de los contribuyentes según su comportamiento.

Por su parte la sub-facturación si bien es una inconsistencia de información para la autoridad fiscal, en este estudio se caracterizó como un comportamiento pues existen ingresos no facturados como la ganancia cambiaria, los rendimientos, el interés acumulable, entre otros que no necesariamente se considera evasión.

En el caso de personas morales el valor crítico que se utilizó para determinar que puede existir sub-declaración de ingresos fue 1.0792 y para personas físicas 1.1531.

Para 2016 las personas morales tienen un porcentaje de ingreso no declarado del 12.6% con una evasión de ISR del 10.2%, en el caso de personas físicas el porcentaje de ingreso no declarado es del 9.8% que representa una evasión del 57.7%, la diferencia en personas físicas entre el porcentaje no declarado y la evasión es por el hecho de que el cálculo de ISR depende de que los rangos de los ingresos y entre más ingreso el porcentaje de ISR es mayor.

Es importante para poder realizar análisis automatizados de una manera más eficiente, rápida, con menos margen de error y con mayor capacidad de análisis que las variables importantes que se capturan al momento de realizar las facturas CFDI por los usuarios no tengan campos libres como por ejemplo el de “Tipo de Moneda”, “Método de Pago” y campo de porcentaje de IVA, para poder así tener catálogos que de manera automática se pueda homogenizar y estandarizar toda la información.

La información de CFDI cuenta con el potencial de realizar fiscalizaciones electrónicas, que disminuyan los costos de la administración tributaria y ampliar el espectro de contribuyentes auditables.

Se debe establecer una figura tipo “ticket-factura” que podría aumentar la capacidad de fiscalización y de disminuir el riesgo de sub-declarar ingresos.

Se recomienda, que de incorporarse el ticket-factura en México, se deben contemplar previamente las dificultades, barreras y malas prácticas que han experimentado otras administraciones tributarias en el mundo.

Referencias

Aggarwal, Charu C (2013), “*Outlier Analysis*”, Springer.

Banxico, Sistema de Información Económica.

<http://www.banxico.org.mx/SieInternet/consultarDirectorioInternetAction.do?sector=6&accion=consultarCuadroAnalitico&idCuadro=CA113&locale=es>

Breiman L., Friedman J. H., Olshen R. A., and Stone, C. J. (1984) “*Classification and Regression Trees*”. Wadsworth.

Fuentes Castro, Hugo Javier. (2016) “*Impacto en la Evasión por la Introducción de la Factura Electrónica*”, ITESM.

INEGI (n.d). *Sistema de Clasificación Industrial de América del Norte 2013 (SCIAN 2013)*. Recuperado de: <http://www.inegi.org.mx/est/contenidos/proyectos/SCIAN/presentacion.aspx>

Morse, S.C., Karlinsky, S., y Bankman, J. (2009). *Cash Businesses and Tax Evasion*. Stanford Law & Policy Review.

OECD (2017). *Technology Tools to Tackle Tax Evasion and Tax Fraud*. Recuperado. <http://www.oecd.org/tax/crime/technology-tools-to-tackle-tax-evasion-and-tax-fraud.htm>

Constitución Política de los Estados Unidos Mexicanos (Reforma 10.06.2011)

Extraída de la [www](http://www.diputados.gob.mx/LeyesBiblio/pdf/1.pdf) el 10 de julio de 2011 de <http://www.diputados.gob.mx/LeyesBiblio/pdf/1.pdf>

Código Fiscal de la Federación (Reforma 10.05.2011)

Recuperada el 10 de julio de 2011 de <http://www.diputados.gob.mx/LeyesBiblio/pdf/8.pdf>

Ley de Impuesto sobre la Renta (Reforma 31.12.2010)

Recuperada el 10 de julio de 2011 de

<http://www.diputados.gob.mx/LeyesBiblio/pdf/82.pdf>

Anexo

Muestra de análisis

Selección de 7 sectores de la economía compuestos por 144 actividades económicas, para el estudio de evasión fiscal.

Sector 46. Comercio al por menor	
Clase	Nombre
461110	Comercio al por menor en tiendas de abarrotes, ultramarinos y misceláneas.
461121	Comercio al por menor de carnes rojas.
461122	Comercio al por menor de carne de aves.
461123	Comercio al por menor de pescados y mariscos.
461130	Comercio al por menor de frutas y verduras frescas.
461140	Comercio al por menor de huevo.
461150	Comercio al por menor de semillas y granos alimenticios, especias y chiles secos.
461160	Comercio al por menor de leche procesada, otros productos lácteos y embutidos.
461170	Comercio al por menor de materias primas para repostería
461190	Comercio al por menor de otros alimentos.
461211	Comercio al por menor de vinos y licores.
461212	Comercio al por menor de cerveza.
461213	Comercio al por menor de refrescos, agua purificada y hielo.
461220	Comercio al por menor de cigarros, puros y tabaco.
462111	Comercio al por menor en supermercados.
462112	Comercio al por menor en minisupers.
462210	Comercio al por menor en tiendas departamentales.
463111	Comercio al por menor de telas.
463112	Comercio al por menor de blancos.
463113	Comercio al por menor de artículos de mercería y bonetería.
463211	Comercio al por menor de ropa, excepto de cuero y piel.
463212	Comercio al por menor de accesorios de vestir.
463213	Comercio al por menor de ropa de cuero y piel y de otros artículos de estos materiales.
463214	Comercio al por menor de pañales desechables.
463215	Comercio al por menor de sombreros.
463310	Comercio al por menor de calzado.

Sector 46. Comercio al por menor	
Clase	Nombre
464111	Farmacias sin minisuper.
464112	Farmacias con minisuper.
464113	Comercio al por menor de productos naturistas y de complementos alimenticios.
464121	Comercio al por menor de lentes.
464122	Comercio al por menor de aparatos ortopédicos.
465111	Comercio al por menor de artículos de perfumería y cosméticos.
465211	Comercio al por menor de discos y casetes.
465212	Comercio al por menor de juguetes y bicicletas.
465213	Comercio al por menor de equipo y material fotográfico.
465214	Comercio al por menor de artículos y aparatos deportivos.
465215	Comercio al por menor de instrumentos musicales.
465311	Comercio al por menor de artículos de papelería.
465312	Comercio al por menor de libros.
465313	Comercio al por menor de periódicos y revistas.
465911	Comercio al por menor de mascotas.
465912	Comercio al por menor de regalos.
465913	Comercio al por menor de artículos religiosos.
465915	Comercio al por menor en tiendas importadoras.
465919	Comercio al por menor de otros artículos de uso personal.
466113	Comercio al por menor de cristalería, loza y utensilios de cocina.
466311	Comercio al por menor de alfombras, cortinas, tapices y similares.
466312	Comercio al por menor de plantas y flores naturales.
466314	Comercio al por menor de lámparas ornamentales y candiles.
466319	Comercio al por menor de otros artículos para la decoración de interiores.
466410	Comercio al por menor de artículos usados.
467111	Comercio al por menor en ferreterías y tlapalerías.
467112	Comercio al por menor de pintura.
467113	Comercio al por menor de vidrios y espejos.
467114	Comercio al por menor de artículos para la limpieza.
467115	Comercio al por menor de materiales para la autoconstrucción.
468211	Comercio al por menor de partes y refacciones nuevas para automóviles, camionetas y camiones.
468212	Comercio al por menor de partes y refacciones usadas para automóviles, camionetas y camiones.

Sector 46. Comercio al por menor	
Clase	Nombre
468213	Comercio al por menor de llantas y cámaras para automóviles, camionetas y camiones.
468411	Comercio al por menor de gasolina y diesel.
468412	Comercio al por menor de gas en cilindros y para tanques estacionarios.
468419	Comercio al por menor de otros combustibles.
468420	Comercio al por menor de aceites y grasas lubricantes, aditivos y similares.

Sector 48-49. Transporte, correos y almacenamientos	
Clase	Nombre
485111	Transporte de pasajeros urbano y suburbano en autobuses.
485113	Transporte de pasajeros urbano y suburbano en trolebuses y trenes ligeros.
485210	Transporte de pasajeros interurbano y rural
487110	Transporte turístico por tierra
487210	Transporte turístico por agua
487990	Otro transporte turístico
488330	Servicios para la navegación por agua.
491110	Servicios postales
492110	Servicios de mensajería y paquetería foránea
492210	Servicios de mensajería y paquetería local

Sector 53. Servicios inmobiliarios y de alquiler de bienes inmuebles	
Clase	Nombre
532110	Alquiler de automóviles sin chofer.
532121	Alquiler de camiones de carga sin chofer.
532122	Alquiler de autobuses, minibuses y remolques sin chofer.
532210	Alquiler de aparatos eléctricos y electrónicos para el hogar
532220	Alquiler de ropa
532230	Alquiler de videocasetes y discos
532291	Alquiler de mesas, sillas, vajillas y similares.
532292	Alquiler de instrumentos musicales.
532299	Alquiler de otros artículos para el hogar.

Sector 53. Servicios inmobiliarios y de alquiler de bienes inmuebles

Clase	Nombre
532310	Centros generales de alquiler
532411	Alquiler de maquinaria para construcción, minería y actividades forestales.
532412	Alquiler de equipo de transporte, excepto terrestre.

Sector 62. Servicios de salud y asistencia social

Clase	Nombre
621311	Consultorios de quiropráctica pertenecientes al sector privado.
621320	Consultorios de optometría
621511	Laboratorios médicos y de diagnóstico pertenecientes al sector privado.

Sector 71. Servicios de esparcimiento culturales y deportivos y recreativos

Clase	Nombre
711111	Compañías de teatro del sector privado.
712111	Museos del sector privado.
712120	Sitios históricos
712131	Jardines botánicos y zoológicos del sector privado.
712190	Grutas, parques naturales y otros sitios del patrimonio nacional
713111	Parques del sector privado con instalaciones recreativas.
713120	Casas de juegos electrónicos
713210	Casinos
713291	Venta de billetes de lotería.
713299	Otros juegos de azar.
713910	Campos de golf
713920	Pistas para esquiar
713930	Marinas turísticas
713941	Clubes deportivos del sector privado.
713943	Balnearios del sector privado.
713950	Boliches
713991	Billares.
713998	Otros servicios recreativos prestados por el sector privado.

Sector 72. Servicios de alojamiento y preparación de alimentos	
Clase	Nombre
722111	Restaurantes-bar con servicio de meseros.
722112	Restaurantes sin bar y con servicio de meseros.
722211	Restaurantes de autoservicio.
722212	Restaurantes de comida para llevar.
722310	Servicios de comedor para empresas e instituciones
722330	Servicios de preparación de alimentos en unidades móviles
722411	Centros nocturnos, discotecas y similares.
722412	Bares y cantinas.
722413	Cervecerías.
722414	Pulquerías.

Sector 81. Otros servicios excepto actividades gubernamentales	
Clase	Nombre
811111	Reparación mecánica en general de automóviles y camiones.
811112	Reparación del sistema eléctrico de automóviles y camiones.
811113	Rectificación de motores de automóviles y camiones.
811114	Reparación de transmisiones de automóviles y camiones.
811115	Reparación de suspensiones de automóviles y camiones.
811116	Alineación y balanceo de automóviles y camiones.
811119	Otras reparaciones mecánicas y eléctricas de automóviles y camiones.
811121	Hojalatería y pintura de automóviles y camiones.
811122	Tapicería de automóviles y camiones.
811129	Instalación de cristales y otras reparaciones a la carrocería de automóviles y camiones.
811191	Reparación menor de llantas.
811192	Lavado y lubricado de automóviles y camiones.
811199	Otros servicios de reparación y mantenimiento de automóviles y camiones.
811211	Reparación y mantenimiento de equipo electrónico de uso doméstico.
811219	Reparación y mantenimiento de otro equipo electrónico y de equipo de precisión.
811313	Reparación y mantenimiento de maquinaria y equipo para mover, acomodar y levantar materiales.

Sector 81. Otros servicios excepto actividades gubernamentales	
Clase	Nombre
811420	Reparación de tapicería de muebles para el hogar
811430	Reparación de calzado y otros artículos de piel y cuero
811491	Cerrajerías.
811492	Reparación y mantenimiento de motocicletas.
811493	Reparación y mantenimiento de bicicletas.
811499	Reparación y mantenimiento de otros artículos para el hogar y personales.
812110	Salones y clínicas de belleza y peluquerías.
812120	Baños públicos.
812130	Sanitarios públicos y boquerías.
812210	Lavanderías y tintorerías.
812410	Estacionamientos y pensiones para automóviles.
812910	Servicios de revelado de fotografías.

Fuente. Elaboración con datos del SAT.

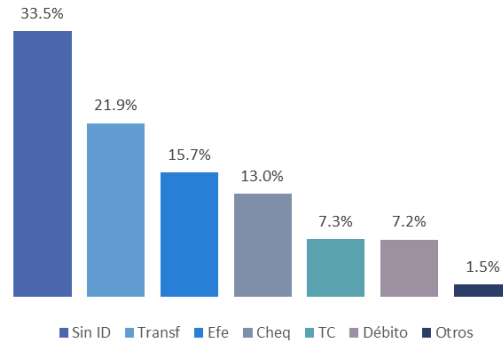
Tratamientos generales de la información

Anonimización: En cuanto a la identificación del contribuyente, el SAT realizó un proceso de pseudonimización de la información, es decir, se sustituyó el RFC por un número entero proporcionado aleatoriamente.

Estandarización: En cuanto la información proveniente del CFDI se observó el método de pago⁵ se captura con texto libre por lo que fue necesario reclasificarlo. El método de pago más común es dejar vacío el campo por lo que se etiquetó como "Sin-ID", le sigue Transferencia, Efectivo, Cheques, Tarjeta de Crédito y Tarjeta de Débito, el resto de métodos de pago se clasificaron como "Otros" por su baja incidencia.

⁵ Forma en que se paga la factura: efectivo, tarjeta de crédito o débito, transferencia bancaria, etc.

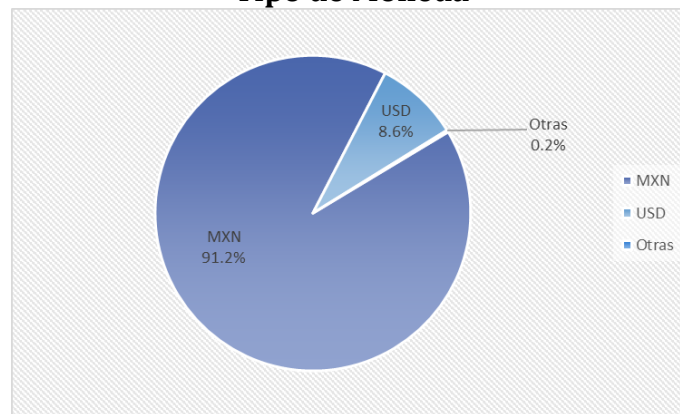
Método de Pago



Fuente. Elaboración con datos del SAT.

En el caso de tipo de moneda se realizó un proceso de reclasificación pues tan solo la moneda nacional contaba con más de 5 denominaciones. Los montos de las facturas con moneda extranjera se multiplicaron por el valor promedio publicado en el Banco de México para tener una estandarización del cambio monetario. Dada la baja incidencia de otras monedas se reclasificaron sólo en moneda nacional “MXN” y extranjera “EXT”.

Tipo de Moneda



Fuente. Elaboración con datos del SAT.

Bases de datos

La base de datos de facturas CFDI emitidas y recibidas tiene la siguiente estructura donde cada registro representa la información contenida en una factura.

Nombre lógico del campo	Relevancia
id_comprobante	Índice
Numero de Operación	Índice
Clave Periodo	Índice
Ejercicio	Índice
Suma de Ingresos Nominales de Mese Anteriores del Ejercicio	Numérico
Ingresos Nominales del Mes que Declara	Numérico
Total de Ingresos Nominales	Numérico
Coefficiente de Utilidad	Numérico
Utilidad Fiscal para Pago Provisional	Numérico
Anticipos y Rendimientos Distribuidos en el Periodo	Numérico
Pérdidas Fiscales de Ejercicios Anteriores	Numérico
Base Gravable del Pago Provisional	Numérico
ISR Causado	Numérico

Fuente. Elaboración con datos del SAT.

Se realizó una primera agregación de la información donde se clasificaron las columnas que son relevantes como tipo índice y las columnas numéricas.

Los criterios que se tomaron para el manejo de la información son los siguientes:

Ejercicio: Año fiscal, 2015 y 2016

Periodo: Mes de la presentación, 1-12.

Moneda: Existe una amplia gama de valores que se captura como texto abierto, se utilizan técnicas de búsquedas de texto para clasificarlas en pesos o monedas de diversos países, en caso de estar vacío se asume como pesos.

Efecto_comprobante: Las facturas son de tipo ingreso, egreso (antes llamadas notas de crédito) y de traslado. Para este estudio se eliminan las de traslado para tener sólo las facturas con efecto fiscal en monto.

RFC_Emisor: ID del RFC que emite la factura y se proporciona de manera anónima.

RFC_Receptor: ID del RFC que recibe la factura y se proporciona de manera anónima.

Método de Pago: Existe una amplia gama de valores que se captura como texto abierto, y se tienen que reclasificar para poder estandarizarlos.

B_ACTIVO: Indica si la factura esta Activa o Cancelada.

Las columnas de carácter numérico se explican con el título de la columna.

Se recibió también un archivo de Declaraciones y Pagos que contiene la siguiente información:

Nombre lógico del campo	Relevancia
Ejercicio	Índice
Periodo	Índice
Moneda utilizada para expresar montos	Índice
Efecto del comprobante fiscal para el contribuyente emisor	Índice
Registro Federal de Contribuyente del Emisor	Índice
Registro Federal de Contribuyente del Receptor	Índice
Metodo de Pago	Índice
banddera de estatus	Índice
Descuento	Numérico
Subtotal	Numérico
Total	Numérico
Total del impuestos trasladados IVA	Numérico
Total del impuestos trasladados IEPS	Numérico
Total del impuestos retenidos IVA	Numérico
Total del impuestos retenidos ISR	Numérico

Fuente. Elaboración con datos del SAT.

De igual manera como en las facturas se realizó una selección de la información relevante para este estudio para determinar las columnas relevantes como tipo índice y las columnas numéricas. Los campos pueden repetirse porque se presentan múltiples declaraciones en un mismo mes y se toma como presentación final la que contenga el número de operación mayor, es decir, la más reciente.

Para relacionar ambas fuentes de información se utilizó la información de id-comprobante, Ejercicio y Periodo, lo que significa que sería en términos prácticos RFC-Año-Mes.

Estandarización de la información

Tipo	Cant	ww	wwAcum	Clasificación
CHEQUE	234,945	2.3%	52.1%	Cheque
Cheque	163,226	1.6%	55.3%	Cheque
2	155,563	1.5%	56.8%	Cheque
CHEQUE NOMINATIVO	99,344	1.0%	66.6%	Cheque
TARJETA DE CREDITO	112,337	1.1%	60.4%	Crédito
4	86,296	0.8%	69.2%	Crédito
TARJETA DE DEBITO	111,832	1.1%	62.5%	Débito

Tipo	Cant	ww	wwAcum	Clasificación
EFFECTIVO	536,835	5.2%	27.2%	Efectivo
Efectivo	500,940	4.8%	32.1%	Efectivo
1	274,419	2.7%	47.4%	Efectivo
99	170,315	1.6%	53.7%	Otros
NO IDENTIFICADO	857,958	8.3%	8.3%	Sin-ID
No identificado	793,705	7.7%	16.0%	Sin-ID
No Identificado	622,308	6.0%	22.0%	Sin-ID
	274,764	2.7%	44.7%	Sin-ID
No aplica	121,530	1.2%	59.3%	Sin-ID
NO APLICA	109,843	1.1%	63.6%	Sin-ID
No Aplica	103,987	1.0%	64.6%	Sin-ID
Desconocido	71,741	0.7%	69.9%	Sin-ID
3	363,641	3.5%	35.6%	Transferencia
TRANSFERENCIA	350,625	3.4%	39.0%	Transferencia
TRANSFERENCIA ELECTRONICA	321,410	3.1%	42.1%	Transferencia
Transferencia	246,525	2.4%	49.8%	Transferencia
Transferencia Electrónica	135,961	1.3%	58.1%	Transferencia
Transferencia electrónica	112,227	1.1%	61.5%	Transferencia
Transferencia Electrónica	103,789	1.0%	65.6%	Transferencia
Transferencia electrónica	92,782	0.9%	67.5%	Transferencia
TRANSFERENCIA BANCARIA	90,102	0.9%	68.3%	Transferencia

Fuente. Elaboración con datos del SAT.

Dado que el método de pago se captura con texto libre es necesario reclasificarlo y se coloca como ejemplo la siguiente tabla que contiene aproximadamente el 70% de los registros.

La nueva clasificación de métodos de pago es la siguiente:

- Sin identificación
- Efectivo
- Cheque nominativo
- Transferencia electrónica
- Tarjeta de crédito
- Monederos electrónicos
- Dinero electrónico
- Tarjetas electrónicas
- Vales de despensa
- Tarjeta de debito
- Tarjeta de servicios
- Otros

En el caso de Moneda se procedió de la misma manera para clasificarlo en la moneda que se realizó la factura.

Texto	Cant	ww	wwAcum	Tipo
MXN	3,288,471	29.68%	29.68%	MXN
NULL	1,986,810	17.93%	47.62%	MXN
Peso Mexicano	1,260,921	11.38%	59.00%	MXN
Pesos	1,025,174	9.25%	68.25%	MXN
MXP	801,643	7.24%	75.49%	MXN
PESOS	652,247	5.89%	81.37%	MXN
MN	401,079	3.62%	84.99%	MXN
M.N.	254,778	2.30%	87.29%	MXN
NACIONAL	116,708	1.05%	90.40%	MXN
Peso	109,532	0.99%	91.39%	MXN
NAL	101,915	0.92%	92.31%	MXN
Pesos Mexicanos	95,045	0.86%	93.17%	MXN
PESOS MEXICANOS	61,688	0.56%	93.72%	MXN
pesos	54,468	0.49%	94.75%	MXN
MONEDA NACIONAL	51,775	0.47%	95.22%	MXN
PESO	47,781	0.43%	95.65%	MXN
USD	227,496	2.05%	89.35%	USD
Dólar Americano	59,611	0.54%	94.26%	USD
DOLARES	37,564	0.34%	95.99%	USD

Fuente. Elaboración con datos del SAT.

Se utilizó información publicada en la página de internet del Banco de México para tener el precio histórico promedio de las monedas y convertir los montos de esas facturas a pesos mexicanos.

Moneda	201501	201502	201503	201504	201505	201506	201507	201508	201509	201510	201511	201512	201601	201602	201603	201604
AUSDOL	11.68	11.71	11.67	12.12	11.78	12.05	11.81	11.90	11.87	11.74	12.01	12.57	12.87	12.93	13.26	13.10
CANDOL	11.82	11.98	12.05	12.69	12.33	12.57	12.36	12.61	12.61	12.56	12.45	12.42	12.93	13.37	13.34	13.72
YUAN	2.40	2.38	2.46	2.48	2.48	2.53	2.59	2.63	2.66	2.61	2.59	2.66	2.77	2.76	2.67	2.65
USD	14.99	14.96	15.26	15.37	15.38	15.69	16.08	16.78	16.91	16.52	16.59	17.25	18.19	18.10	17.24	17.18
LIBRA	22.51	23.11	22.66	23.62	23.47	24.67	25.09	25.81	25.61	25.43	24.96	25.44	25.82	25.23	24.78	25.15
YEN	0.13	0.13	0.13	0.13	0.12	0.13	0.13	0.14	0.14	0.14	0.13	0.14	0.15	0.16	0.15	0.16
FRANCO_SWISS	16.29	15.77	15.71	16.40	16.30	16.77	16.72	17.36	17.30	16.75	16.13	17.25	17.77	18.16	18.00	17.92
EURO	16.91	16.77	16.40	17.22	16.86	17.48	17.76	18.81	18.87	18.26	17.52	18.75	19.69	19.67	19.64	19.67

Fuente. Elaboración con datos del SAT.

Moneda	201605	201606	201607	201608	201609	201610	201611	201612	201701	201702	201703	201704	201705	201706	201707
AUSDOL	13.34	13.75	14.27	14.17	14.83	14.34	15.16	14.93	15.79	15.37	14.34	14.17	13.91	13.85	14.26
CANDOL	14.08	14.22	14.39	14.37	14.75	14.08	15.26	15.37	15.94	15.10	14.10	13.87	13.84	13.91	14.28
YUAN	2.80	2.78	2.83	2.82	2.91	2.79	2.97	2.97	3.03	2.91	2.73	2.75	2.74	2.67	2.65

Moneda	201605	201606	201607	201608	201609	201610	201611	201612	201701	201702	201703	201704	201705	201706	201707
USD	18.41	18.46	18.78	18.86	19.38	18.89	20.52	20.62	20.79	20.00	18.80	18.96	18.69	18.06	17.86
LIBRA	26.80	24.78	24.92	24.71	25.16	22.96	25.61	25.48	26.16	24.87	23.49	24.52	24.12	23.45	23.54
YEN	0.17	0.18	0.18	0.18	0.19	0.18	0.18	0.18	0.18	0.18	0.17	0.17	0.17	0.16	0.16
FRANCO_SWISS	18.53	18.96	19.41	19.19	19.98	19.08	20.14	20.29	21.07	19.96	18.77	19.05	19.31	18.86	18.50
EURO	20.51	20.52	20.99	21.00	21.77	20.67	21.76	21.75	22.46	21.24	20.10	20.65	21.01	20.60	21.06

Fuente. Elaboración con datos del SAT.

Se recomienda ampliamente que el SAT establezca un catálogo único de valores permitidos en ambos campos de Método de Pago y Tipo de Moneda, para que en análisis automatizados de la información sea mucho más eficiente el uso de esta información.

Cálculo de ISR

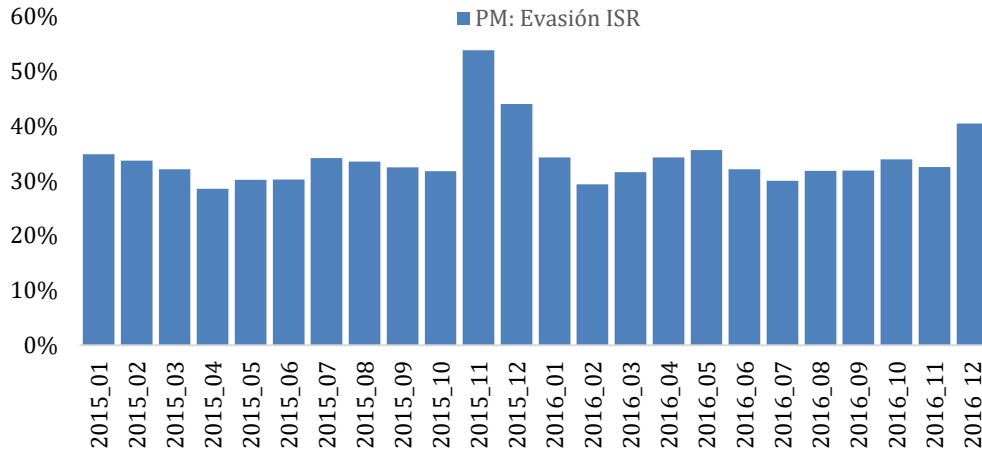
El cálculo de ISR para personas morales se calcula a partir de los ingresos mensuales en su declaración mensual multiplicado por el factor de utilidad del año que le corresponda y por la tasa de impuesto del 30%. Para el cálculo de ISR potencial se utiliza como ingresos mensuales el total de las facturas CFDI generadas.

En el caso de personas físicas, se toma como ingreso base el total de ingresos de su declaración mensual restando los gastos declarados en el mes, posteriormente este ingreso base se determina en que rango de la tabla de la LISR del artículo 113, para aplicar entonces la cuota fija más el porcentaje del excedente del límite inferior. Para el cálculo del ISR potencial se utiliza como ingresos mensuales el total de facturas CFDI generadas y se repite el procedimiento anterior.

Impuesto - Artículo 113			
Límite inferior	Límite superior	Cuota fija	% Excedente del lím. inferior
0.01	496.07	0.00	1.9%
496.08	4210.41	9.52	6.4%
4210.42	7399.42	247.24	10.9%
7399.43	8601.50	594.21	16.0%
8601.51	10298.35	786.54	17.9%
10298.36	20770.29	1090.61	21.4%
20770.30	32736.83	3327.42	23.5%
32736.84	62500.00	6141.95	30.0%
62500.01	83333.33	15070.90	32.0%
83333.34	250000.00	21737.57	34.0%
250000.01	... y más	78404.23	35.0%

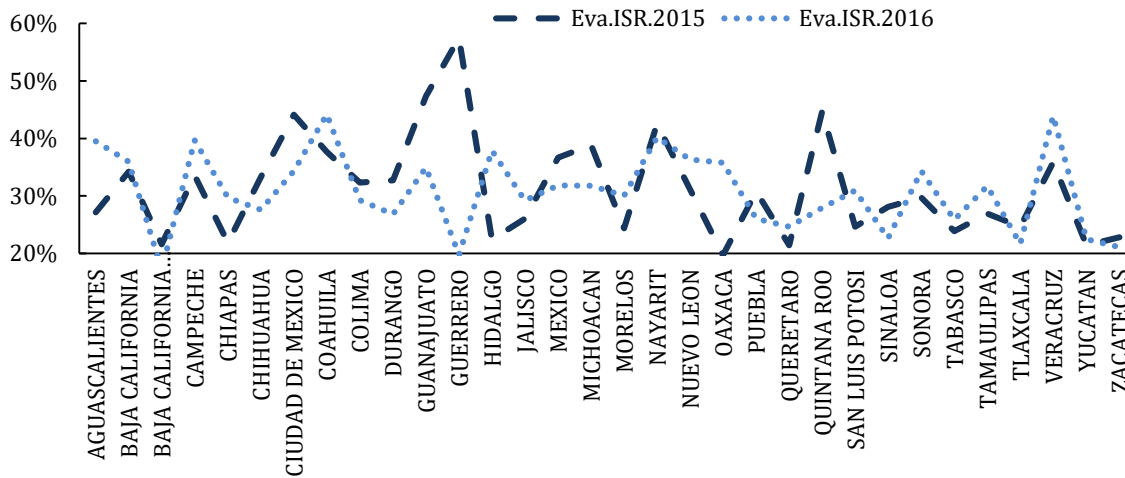
Tasas de evasión de contribuyentes que evaden impuestos

Los contribuyentes registrados como personas morales que tienen indicios de estar sub-declarando ingresos tienen unas tasas de evasión de 37.1% para el 2015 y del 34.0% para el 2016. El comportamiento mensual nos indica que existe tendencia a sub-declarar más en los meses de noviembre y diciembre.



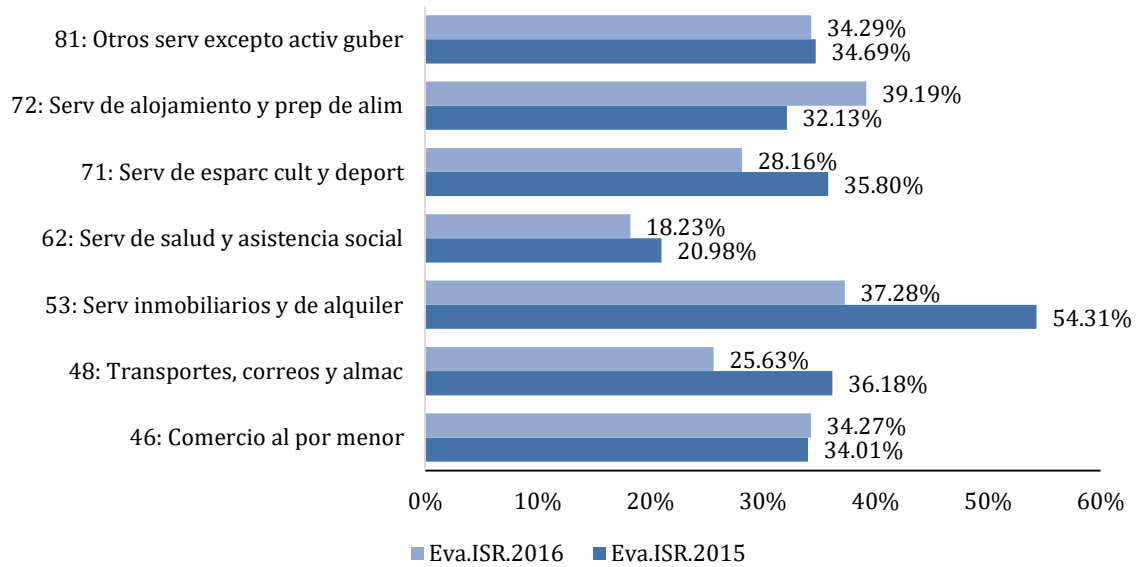
Fuente. Elaboración con datos del SAT.

En los estados se tiene que los estados de Guerrero y Quintana Roo tenían una tasa muy fuerte en el 2015 pero con una disminución importante para el 2016.



Fuente. Elaboración con datos del SAT.

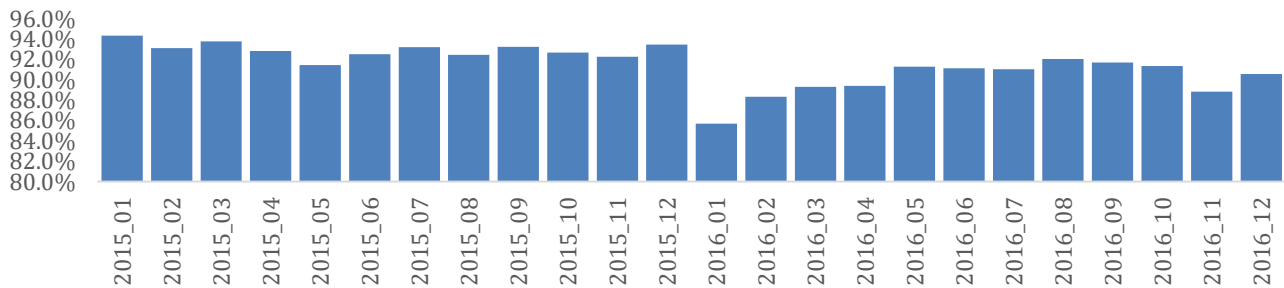
Por sector destaca el sector Inmobiliario y de alquiler en el 2015 con una mejoría importante para el 2016.



Fuente. Elaboración con datos del SAT.

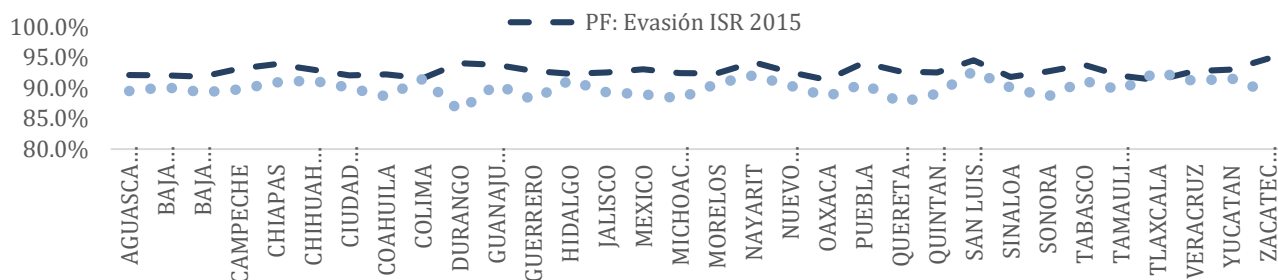
Los contribuyentes registrados como personas físicas que tienen indicios de estar sub-declarando ingresos tienen unas tasas de evasión del 93.4% para el 2015 y del 89.5% para el 2016. El comportamiento mensual nos indica que en el 2016 que hay una disminución generalizada.

PF: Evasión ISR



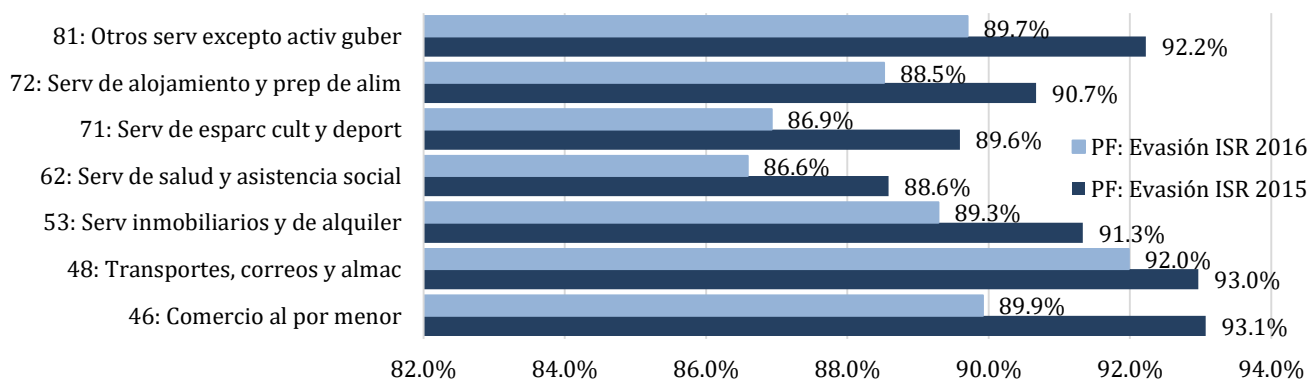
Fuente. Elaboración con datos del SAT.

En los estados también tienen una disminución generalizada en el año 2016 en comparación con el año 2015.



Fuente. Elaboración con datos del SAT.

Por sector se tiene de igual manera una disminución generalizada del año 2016 en comparación del año 2015, destacando el sector Comercio al por menor como el más alto en el año 2015 y que sigue siendo el segundo más alto en el 2016 sólo por debajo del sector Transportes, correos y almacenamiento.



Fuente. Elaboración con datos del SAT.

Análisis de aprendizaje automático

Con el objetivo de encontrar patrones de comportamiento diferenciados entre las empresas que pueden ser declaradas normales y las que tienen indicios de sub-declaración se construyeron las siguientes variables con la información disponible para facturas emitidas y recibidas:

- Monto de facturas canceladas.
- Monto de facturas tipo egreso.
- Monto de Descuento.
- Monto de impuestos trasladados de IVA.
- Monto de impuestos trasladados de IEPS.
- Monto de impuestos retenidos de IVA.
- Monto de impuestos retenidos de ISR.
- Monto pagado por tipo de pago. (Sin-ID, efectivo, transferencia, cheque, etc)
- Monto pagado por tipo de moneda. (Moneda nacional y extranjera)

Estos montos fueron transformados a montos relativos por el monto de facturas de ingreso y activas de facturas emitidas y recibidas respectivamente. Para con esto tener sólo como dato representativo de cada empresa su comportamiento en términos relativos y poder así hacer comparable una empresa chica con una empresa grande.

Sólo en el caso de Monto de facturas recibidas y activas se realizó el cálculo dividiendo esta cantidad entre el Monto de facturas emitidas y activas. De esta manera se tienen entonces una relación de facturas que emite en comparación de las que recibe.

Realizando una prueba de hipótesis estadística de igualdad de medias para cada una de las variables de empresas declaradas como normales en comparación de las que se tiene indicios de sub-declaración se tienen las siguientes variables donde la diferencia es estadísticamente significativa al 5%.

Variable	Prom NORMAL	Prom SUB-DECL
Facturas tipo egreso	2.5%	5.4%
Descuento en facturas	0.6%	5.2%
Sin identificación del pago	64.1%	69.9%
<u>Relación Facturas Recibidas/ Emitidas</u>	<u>99.7%</u>	<u>85.0%</u>

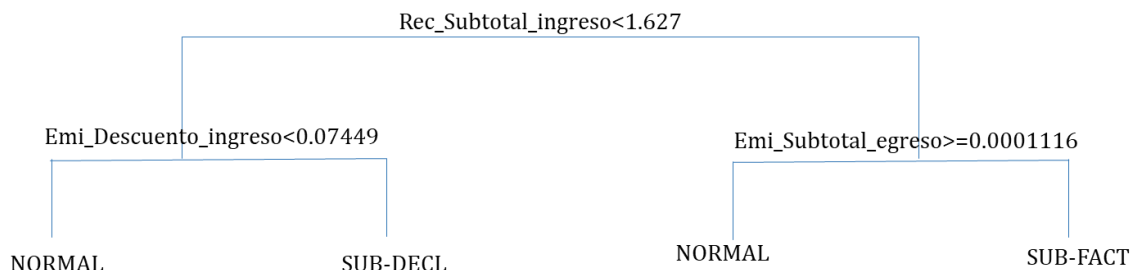
Fuente. Elaboración con datos del SAT.

La tabla anterior nos indica que:

- Las empresas que tienen sub-decl tienen más facturas de tipo egreso (notas de crédito).
- Las empresas que tienen sub-decl emiten más descuentos en sus facturas de ingresos.
- Las empresas que tienen sub-decl tienen una mayor de proporción de pagos sin identificar.

- Las empresas que tienen sub-fact tienen una relación de 85% de facturas recibidas/facturas emitidas mientras que las empresas normales es del 99.7%.

Construyendo un árbol de clasificación mediante la técnica de partición recursiva y regresión con las mismas variables para el diagnóstico descriptivo anterior, se tiene el siguiente árbol de decisión.



Fuente. Elaboración con datos del SAT.

Estos árboles generan las mejores particiones de las variables con el objetivo de hacer la mejor clasificación de los datos.

La primera variable nos indica que variable es la que discrimina mejor las diferentes clasificaciones de los datos, que en este caso es la relación del monto de facturas recibidas sobre monto de facturas emitidas, y así en orden descendente del árbol.

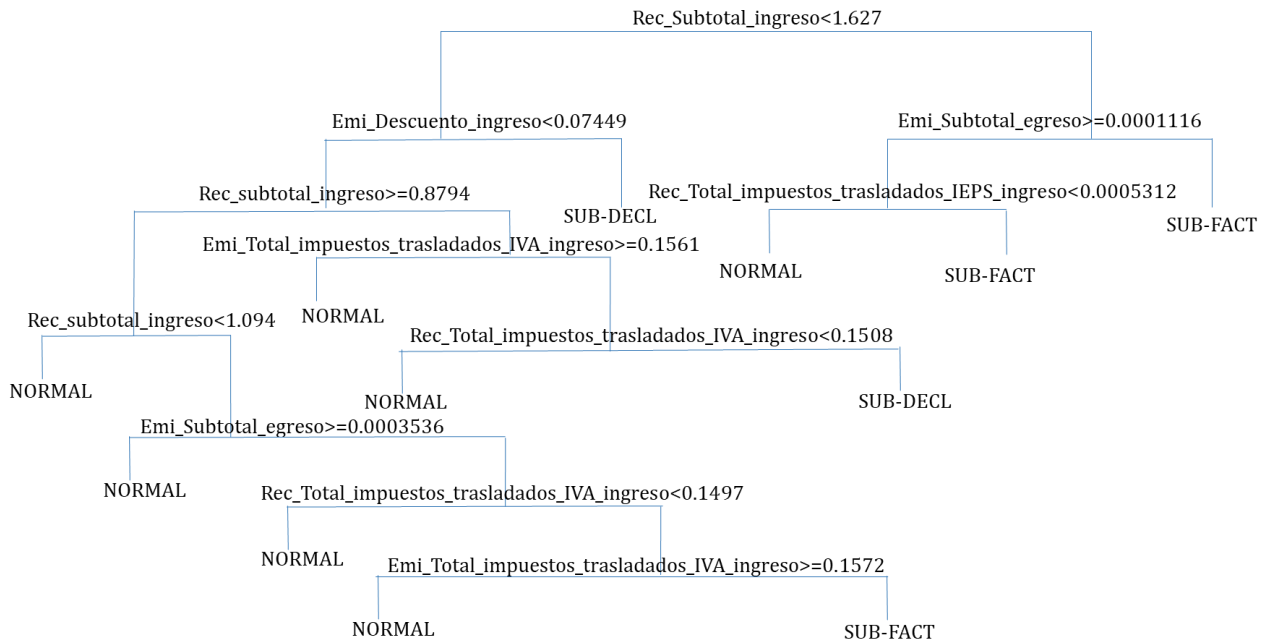
Posteriormente revisamos la condición que está planteada en el encabezado de la rama y si se cumple nos seguimos a la siguiente condición por la izquierda y si no por la derecha.

Si la rama no continua entonces se llegó a una clasificación estimada.

Por ejemplo, el caso para llegar una clasificación de SUB-DECL, se tiene que cumplir que

- La relación entre monto de facturas recibidas y monto de facturas emitidas sea menor a 1.517.
- Este conjunto de empresas que cumplen esta condición se busca una segunda condición que el monto de descuento en las facturas emitidas sea mayor a 0.07799.
- Las empresas que cumplen conjuntamente las 2 condiciones anteriores tienen una probabilidad de $(4363)/(4363+1087+169) = 77.6\%$

Aumentando el criterio de profundidad en la generación de las variables de decisión se genera el siguiente árbol de decisiones:



Fuente. Elaboración con datos del SAT.

Donde las ramas principales siempre son las mismas, sólo que el detalle de las clasificaciones es más fino.