

METODOLOGÍA – Estimación de la tasa efectiva de impuesto sobre la renta por nivel de ingreso.

Este documento se divide en cuatro partes: (I) objetivo del ejercicio, (II) metodología, (III) limitaciones de los datos (IV) resultados.

I. Objetivo

Determinar qué porcentaje de su ingreso dedican los mexicanos al pago del ISR según su nivel de ingreso.

II. Metodología

1. Medición de los ingresos y el pago de ISR. Los datos utilizados en la medición provienen de las bases de datos de declaraciones anuales publicadas por el Sistema de Administración Tributaria (SAT) en 2015. Para este ejercicio se utilizan dos tipos de bases de datos distintas correspondientes a información de personas físicas.

En primer lugar, está la *Declaración Informativa Múltiple de sueldos y salarios* (DIM). Esta base de datos se publica para los ejercicios fiscales de 2013, 2014 y 2015. Incluye los datos de las declaraciones anuales de todos los trabajadores asalariados formales. El ingreso se mide con la variable “i_c1_300101_AA” (Total de Sueldos, Salarios y Conceptos Asimilados).

En segundo lugar está la base de datos de *Declaraciones Anuales de Personas Físicas (anonimizada)* (PF). Ésta se publicó para los ejercicios fiscales de 2010 a 2015, pero los ejercicios de 2010 a 2012 no se utilizan ya que no se cuenta con la información proveída por la DIM para estos años. Esta base de datos mide los ingresos de las personas físicas con ingresos formales pero que no reciben salarios, sino otro tipo de remuneraciones como honorarios, rentas, ingreso por ventas, etc. Para medir el ingreso se usa la variable “I_DEC_TIAONCT1_AA” (Total de Ingresos Acumulables).

Los RFCs en las bases de datos fueron alterados por el SAT para mantener la anonimidad de las personas, pero el proceso de anonimización se realizó de tal forma que el RFC alterado es el mismo en todas las bases de datos y años en las que está presente el RFC original. Debido a que las personas pueden tener ingresos tanto por salarios como por otros medios, además de distintos trabajos en un mismo año, es posible que un mismo RFC aparezca en dos bases de datos del mismo año, o inclusive más de una vez dentro de la misma base de datos.

Para obtener los totales de ingreso e ISR de cada persona, primero se pegaron las bases de datos DIM y PF del mismo año y se realizó una sumatoria de las variables de ingreso e ISR por RFC anonimizado. Con esto se obtiene una base de datos anual con una sola observación para cada ISR, pero que ya incluye el total de ingresos e ISR acumulados a partir de sus diferentes fuentes de ingreso.

Debido a que no todas las personas declaran todos los años o pueden tener ingresos muy distintos de un año a otro, se promedian los ingresos y pagos de ISR de los tres años disponibles para tener una mejor idea de lo que regularmente ganan y pagan las personas. Solo se promedian los años en los que aparece el ISR. Es decir, si un RFC aparece solamente en 2015 se reportan los valores de 2015 tal cual, y si aparece en 2014 y 2015 se promedian esos dos años, no se calcula el promedio como si los años no disponibles tuvieran valores de 0.

Se elimina cualquier observación que, después de calcular el promedio, tenga un ingreso de 0.

2. Cálculo de las tasas efectivas por nivel de ingreso. Con los promedios de ingreso e ISR para todos los años y personas se divide el ISR entre el ingreso para obtener la tasa efectiva. Es decir, el porcentaje del ingreso que realmente se dedica al pago de ISR, en contraste con lo que establece la ley.

Para obtener los niveles de ingreso, se ordenan todas las observaciones de menor a mayor y se dividen en deciles de ingresos. El decil X se divide nuevamente en 10 para obtener los percentiles 91 a 100. Nuevamente el percentil 100 se divide en 10 y este proceso se repite para cada nivel de ingreso más alto hasta obtener el 0.001% más rico de la población.

Los niveles de ingreso se agregan de la siguiente forma:

- Los deciles I a IX se reportan sin cambios
- El siguiente nivel corresponde al decil X menos el 10% más rico dentro del decil
- El siguiente corresponde al 1% más rico de la población menos el 10% más rico dentro del 1%
- El proceso se repite hasta llegar al 0.001% más rico que se reporta completo.

Una vez que los niveles de ingreso se asignaron se calculan los promedios de ingreso, ISR y tasa efectiva por nivel.

3. Determinantes de la tasa efectiva. Para entender el efecto que tiene cada variable sobre la tasa efectiva de ISR se realizó una regresión lineal de mínimos cuadrados ordinarios (OLS) con errores robustos. La variable dependiente es la tasa efectiva, y las variables explicativas son el ingreso y una variable binaria con valor de 1 si la observación se encuentra dentro del 0.1% más rico de la población y 0 en cualquier otro caso. Esto debe ilustrar si, como regla general, mayores ingresos implican mayores tasas efectivas, y cuál es el efecto particular de pertenecer al 0.1% más rico de la población. La regresión se puede resumir en la siguiente ecuación, en donde cada β representa el coeficiente de regresión de las variables explicativas:

$$TasaEfectiva = \beta_1 \cdot Ingreso + \beta_2 \cdot 0.1pct$$

Coeficientes negativos indicarían que un aumento en la variable reduce la tasa efectiva mientras que coeficientes positivos indican lo opuesto

III. Limitaciones

Los cálculos tienen dos limitantes principales: (a) la alteración de los datos para mantener la anonimidad y (b) la exclusión de los ingresos no reportados.

A. Alteración de los datos. Además de alterar los RFCs, el SAT altera los valores de ingresos y de pago de ISR para un mayor nivel de anonimidad. Esto quiere decir que los cálculos no son exactos. Sin embargo, de acuerdo con la explicación proveída por el SAT mismo, esto se hace de tal manera que las propiedades de cada variable como la media y la desviación estándar se mantengan constantes, así como las correlaciones entre las diferentes variables. El proceso de alteración de los valores se hace sumando un “ruido” a cada valor individual de tal forma que se oculte el valor original. Estos ruidos son números aleatorios con media de 0 y desviación estándar seleccionada para conservar las propiedades estadísticas.

Si bien, esto no permite sacar conclusiones a nivel de observación, los resultados agregados deben de ser representativos de lo que se observa en la realidad, de acuerdo con la metodología del SAT.

B. Ingresos no reportados. Los ingresos informales, por definición, no se declaran ante el SAT. Adicionalmente, Otros mecanismos como la elusión y la evasión fiscal, permiten a las personas esconder algunos de sus ingresos por diferentes métodos. Esta es una limitante que no puede resolverse en la ausencia de otro método para determinar los ingresos de las personas. Por lo tanto, los resultados deben tomarse con ello en mente, y bajo el supuesto de que las tasas efectivas pueden estar sobreestimadas para la población general y que, como se presentan, son válidos únicamente para trabajadores formales.

IV. Resultados

En la Tabla 1 se presentan los promedios de ingreso, ISR y tasa efectiva por niveles de ingreso. Como se puede ver, la tasa aumenta conforme aumenta el nivel de ingreso, excepto para los niveles más altos. Los ingresos se presentan en términos mensuales y deflactados a precios de 2021.

Tabla 1. Promedio de ingreso, ISR y tasa efectiva por nivel de ingreso

NIVEL DE INGRESO	INGRESO	TASA EFECTIVA
I	\$ 206.64	0.6%
II	\$ 839.85	0.6%
III	\$ 1,878.54	0.8%
IV	\$ 2,977.62	0.9%
V	\$ 4,185.66	1.4%
VI	\$ 6,336.62	2.0%
VII	\$ 9,780.11	3.7%
VIII	\$ 15,477.98	6.3%
IX	\$ 25,166.62	8.6%
TOP-10% (SIN TOP-1%)	\$ 57,150.01	10.0%
TOP-1% (SIN TOP-0.1%)	\$ 187,496.77	16.0%
TOP-0.1% (SIN TOP-0.01%)	\$ 634,527.75	20.9%
TOP-0.01% (SIN TOP-0.001%)	\$ 1,815,947.47	20.1%
TOP-0.001%	\$ 5,968,662.58	11.3%

La disminución en los ingresos más altos también se confirma en los resultados de la regresión que se pueden ver en la Tabla 2.

Tabla 2. Determinantes de la tasa efectiva de ISR.

TASA EFECTIVA	COEF.	ROBUST STD. ERR.	T	P>T	[95% CONF. INTERVAL]
INGRESO***	4.75E-06	4.27E-07	11.11	0.000	3.91E-06 5.58E-06
0.1PCT***	-19.1	2.82	-6.76	0.000	-24.6 -13.6
_CONS	4.10	0.37	11.0	0.000	3.36 4.83
N = 45,859,330					

Tanto el coeficiente del ingreso como el de la variable binaria de pertenecer al 0.1% más rico son significativos con un nivel de confianza de 99%. Sin embargo, el primero de estos es positivo indicando que mayor ingresos implican mayores tasas efectivas en general, mientras que el segundo es negativo, indicando que pertenecer al 0.1% más rico de la población implica menores tasas efectivas.