

DETERMINANTES DE LA DIFERENCIACIÓN SALARIAL EN MÉXICO

Determiners of the wage breach in Mexico

JÉSSICA GÁMEZ ARROYO - jgameza@uaemex.mx

EDUARDO ROSAS ROJAS - erosasr@uaemex.mx

Resumen

En este artículo se analizan los determinantes de la diferenciación salarial en México, para lo cual se incluyen características de los trabajadores, tales como: edad, sexo, años de escolaridad y estado conyugal; variables de la empresa en donde labora, tales como: tamaño de la empresa, prestaciones laborales, si existe sindicato y rama; y variables regionales, tales como la ciudad en la que vive el trabajador y el tamaño de la localidad. Dichas variables son obtenidas para el primer trimestre de 2015 de la ENOE (Encuesta Nacional de Ocupación y Empleo). Se aplica la técnica de modelos econométricos multinivel, las variables significativas que dan cuenta de la diferenciación salarial son la edad, el tamaño de la empresa, prestaciones laborales y la ciudad. Los resultados obtenidos son contrarios a los postulados de la teoría del capital humano, la cual asegura que la escolaridad es la causa principal de las diferencias en los ingresos de los trabajadores.

Palabras clave: México; diferenciación salarial; modelos econométricos; capital humano; ingresos.

Abstract

This article analyzes the determiners of the wage breach in Mexico, which include employees' characteristics such as: age, genre, education level and marital status. Variables of the company like its scope, job benefits, the existence of a union or branch and regional variables as the city of residence and the extent of the town are also explained. These variables are obtained through the ENOE (National Survey of Occupation and Employment, for its initials in Spanish). The multi-level econometric model technique is applied to learn that the significant variables that generate the wage breach in Mexico are age, size of the enterprise, job benefits and the city. The obtained

results are opposite to the postulates in the theory of human capital which assure that the education level is the principal cause for the income difference amongst workers.

Keywords: Mexico; wage breach; econometric models; human capital; income.



Marco Teórico

Cue y Quintana (2009) señalan que “el trabajo consiste en la capacidad y habilidad del ser humano para producir bienes y servicios” (p.201). En el análisis económico se considera que el salario equilibra el mercado de trabajo, es decir, es la variable que hace que la oferta de trabajo y demanda de trabajo se igualen.

En el artículo 82 de la Ley Federal de Trabajo, vigente en 2015, se establece que “salario es la retribución que debe pagar el patrón al trabajador por su trabajo”, mientras que en el artículo 83 se menciona que “el salario puede fijarse por unidad de tiempo, por unidad de obra, por comisión, a precio alzado o de cualquier otra manera” (p.20).

Los orígenes del actual concepto de capital humano se pueden remontar hasta las obras de los autores clásicos, quienes pusieron de manifiesto que determinadas cualidades humanas, tales como los conocimientos o habilidades adquiridas por los trabajadores, su salud, virtudes, etc., influyen en la actividad económica.

La teoría del capital humano parte de los supuestos del enfoque neoclásico, tales como la competencia perfecta; en su análisis se incorpora la idea de considerar al trabajador como capital. Los trabajos pioneros de Schultz (1961) y Becker (1964) postulaban que las personas invertían en sí mismas para adquirir conocimientos que eran susceptibles de ser valorados económicamente, y que podían ser asimilados a un tipo de capital cuyo rendimiento podía calcularse mediante las ecuaciones de Mincer (1974). De esta forma, la educación formal en la que invertían los agentes económicos pasó a constituir un pilar básico de la teoría del capital humano.

Por lo tanto, dicha teoría explica que las personas suelen dedicar mayor o menor tiempo y recursos para adquirir cualificaciones o educación; de esta forma, los trabajadores llegan al mercado de trabajo con diferentes calificaciones, que dependen en primera instancia, de las cualidades innatas de la persona

y, en segundo lugar, del tiempo que ha dedicado en adquirirlas, es decir, del tiempo que ha dedicado a invertir en capital humano. Estas capacidades pueden ser adquiridas en distintos ámbitos: a través de la educación -ya sea formal o informal-, en la escuela o familia, o simplemente a través de la experiencia.

Schultz (1961) argumenta que los recursos humanos tienen dimensiones cuantitativas y cualitativas; de manera particular, él se enfoca en el estudio de las segundas. Dentro de los componentes cualitativos identifica la habilidad, los conocimientos y atributos para realizar el trabajo productivo (p.1).

En palabras de Schultz (1985) “propongo tratar la educación como una inversión en el hombre y tratar sus consecuencias como una forma de capital. Como la educación viene a formar parte de la persona que la recibe, me referiré a ella como capital humano”.

Gary Becker (1964) y Jacob Mincer (1974) formularon el modelo matemático, conocido como la ecuación minceriana, en donde se relacionan los salarios con las inversiones en capital humano. Este enfoque le transfiere de forma individual y exclusiva al trabajador la responsabilidad de su nivel de ingreso. La ecuación básica de Mincer (1974, p.116) es:

$$\ln y_i = \alpha_0 + \rho_s S + \beta_0 X + \beta_1 X^2 + \varepsilon$$

Donde $\ln y_i$ es el logaritmo natural de los ingresos, ρ_s es la tasa de retorno a la educación S , β_0 es el coeficiente que mide el efecto de la experiencia (X) sobre los ingresos y β_1 trata de captar la concavidad de los perfiles de ingreso con respecto a la experiencia.

El valor del coeficiente de los años de educación formal se interpreta como la tasa de rendimiento media de un año adicional de estudio de los trabajadores. Por otro lado, conforme aumenta la experiencia, los ingresos individuales aumentan, pero cada año de experiencia tiene un efecto sobre los ingresos menor que el anterior, por lo que se espera que al ser la función cóncava con relación a la experiencia, la estimación de β_0 sea positiva y la de β_1 sea negativa.

Becker (1964) concluye que la mayor parte de las inversiones en capital humano -como la educación formal, la formación en el trabajo, o las migraciones- elevan las retribuciones a edades avanzadas. Además de que si algunos individuos ganan más que otros, es porque invierten más en sí mismos (pp. 29-33).

Garza (2013) estimó un modelo para México que incluye las habilidades en el puesto de trabajo bajo un enfoque de regionalización. Concluye que la relación salarial en las regiones es un fenómeno multifactorial en donde no sólo la escolaridad es determinante, sino que, aún con mayor robustez, influyen la localización, el tamaño de la empresa, la rama de actividad económica, y, sobre todo, las habilidades que el trabajador desarrolla en su ocupación específica (pp. 174-175). Destaca que, en un contexto de compresión salarial generalizado en un período de estudio –que va del 2005 al 2010–, el bono salarial de los trabajadores sindicalizados mexicanos creció, en el sentido de que la existencia de un sindicalismo fuerte parece reducir la desigualdad salarial y que la acción de los sindicatos es determinante en la protección del salario y la búsqueda de mejores condiciones de trabajo (p.173). Dicha investigación evidencia que la teoría del capital humano no se cumple en el mercado laboral mexicano.

Se ha encontrado evidencia de que el tamaño de la empresa es una variable que influye en la determinación del nivel de ingresos (Rojas, Angulo y Velázquez, 2000, p. 131; Garza, 2014, p. 174); encontrando una relación positiva entre el tamaño de la empresa y los ingresos, es decir, se perciben mayores salarios en las empresas de mayor tamaño. En estas investigaciones se mide el tamaño de la empresa por el número de empleados.

La justificación que se da para este hecho es que las condiciones laborales son diferentes en las empresas grandes respecto a las pequeñas; esto debido a que las empresas grandes tienden a ser más intensivas en capital, demandan trabajadores que posean más conocimientos y habilidades que complementen dicho capital.

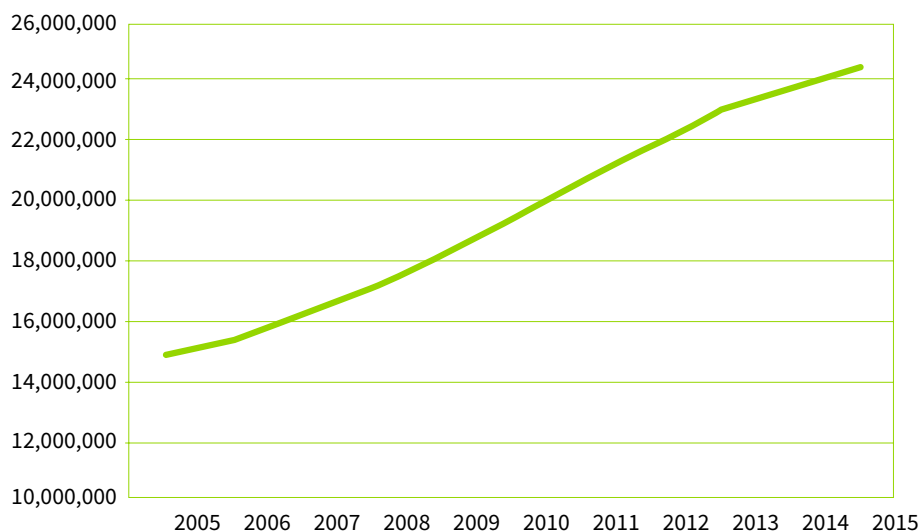
Se ha tratado de justificar que los trabajadores de las grandes empresas son los más beneficiados, debido a la existencia de estructuras sindicales y contratos más fuertes, más propios de las empresas grandes, y que si se convirtieran en trabajadores de empresas pequeñas perderían dichos beneficios. La existencia de los contratos laborales afecta más a los trabajadores menos cualificados que a los de alta dirección o especialización.

Análisis estadístico de los datos

La teoría del capital humano asegura que para disminuir las brechas salariales es necesario invertir en educación; sin embargo, la evidencia empírica para México muestra que los años de escolaridad de la población han aumentado significativamente, y esto no ha ido acompañado de una reducción en las

brechas salariales. En la Gráfica 1 se muestra que las personas con nivel de instrucción medio superior y superior han aumentado de 2005 a 2015.

Gráfica 1. Nivel de instrucción: medio superior y superior



Fuente: Elaboración propia con datos de la ENOE, INEGI.

Para el presente análisis se toman en cuenta los datos de la ENOE correspondientes al primer trimestre de 2015; se consideran sólo a los individuos que perciben un salario por su trabajo y se eliminaron de la base de datos a aquellos que respondieron NS (No sabe) en alguna de las preguntas de interés. La muestra es de 23,875 personas, que al aplicar el factor de expansión, representan a 5,382,748 trabajadores.

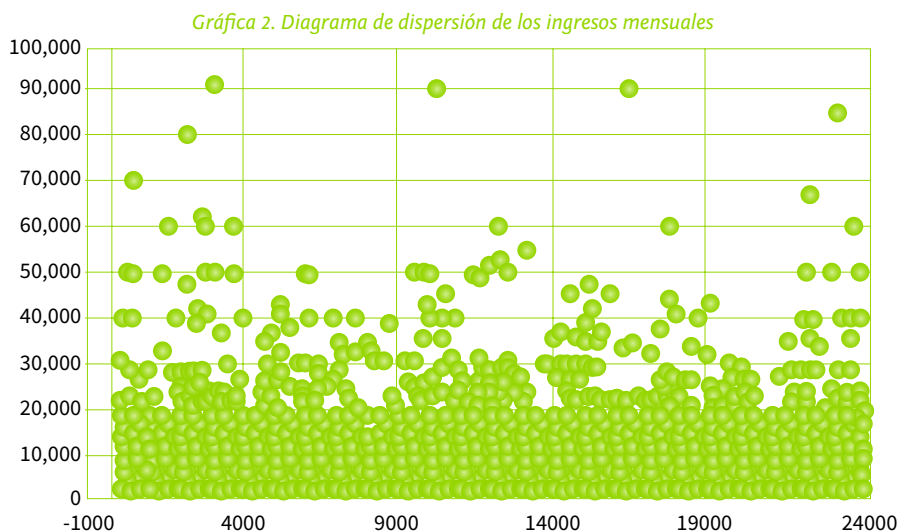
La mayoría de las variables se recodificaron como dicotómicas, construidas de la siguiente forma (Tabla 1):

Tabla 1. Descripción de variables

Clasificación de variables	Variable	Significado
Características del trabajador	Edad	Es la edad del entrevistado, se considera de los 12 a los 98 años.
	Sexo	1 = si es hombre y 0 = si es mujer.
	Años de escolaridad	Son los años cursados en la escuela que reporta el individuo, los códigos van de 0 a 24.
	Estado conyugal	1= si el individuo es casado o vive en unión libre, 0 = cualquier otro estado civil.
Características del trabajo	Sindicato	1 = cuando el trabajador pertenece a algún sindicato, y 0= en caso contrario.
	Contrato	1= cuando tiene contrato laboral escrito , y 0 en cualquier otro caso.
	Aguinaldo	1= si el trabajador reporta percibir aguinaldo, y 0= si no recibe aguinaldo.
	Vacaciones	1=tiene vacaciones con goce de sueldo y 0 = no tiene vacaciones con goce de sueldo.
	Utilidades	Es una variable dummy (1=recibe utilidades y 0=no recibe utilidades).
	Crédito a la vivienda	Es una variable dummy (1=tiene crédito para la vivienda y 0=no tiene crédito para la vivienda)
	Tamaño de la empresa	Se construyeron cuatro variables dummy: MICRO: 1 si laboran de 1 a 10 personas, 0 en cualquier otro caso. PEQUEÑA: 1 si laboran de 11 a 50 personas, 0 en cualquier otro caso. MEDIANA: 1 si laboran de 51 a 250 personas, 0 en cualquier otro caso. GRANDE: 1 si laboran de 251 a más personas, 0 en cualquier otro caso.
	Ingreso mensual	Es el ingreso mensual promedio (en pesos) del trabajador.
	Rama	Se construyó una variable binaria para cada rama, asignando el valor de 1 cuando el trabajador labore en la rama mencionada y 0 en cualquier otro caso. Las rama a considerar son: agropecuario, industria manufacturera, construcción, comercio y servicios.
	Seguridad social	Se construyó una variable dummy, en la cual 1= si tiene acceso a atención médica en cualquier institución, 0= si no tiene acceso a atención médica.
Características del lugar de residencia del trabajador	Ciudad	Es el identificador para cada una de las ciudades, es de dos dígitos que van del 01 al 43.
	Tamaño de la localidad	Se le asigna el valor de 1= si en la localidad hay más de 100,000 habitantes y 0 = en cualquier otro caso.

Fuente: Elaboración propia

En la Gráfica 2 se presenta un diagrama de dispersión de los ingresos mensuales de los trabajadores asalariados, y resulta evidente que existe una gran variación en los datos. La mayor parte de los asalariados obtiene ingresos inferiores a \$10,000 mensuales, aunque existen individuos que reportan ingresos de casi \$100,000 mensuales, lo cual resalta la heterogeneidad de los salarios en México.



Fuente: Elaboración propia con datos de la ENOE, INEGI.

Existen diferencias salariales dependiendo de ciertas características que posea el trabajador, así como de las particularidades de su trabajo y del lugar en el que vive. En la Tabla 2 se registran los datos correspondientes a promedios de ingresos mensuales. La media total de ingresos mensuales de los asalariados es de \$6275.42; este dato no dice mucho por sí solo, pero cuando se consideran los datos por subgrupos resaltan las brechas salariales.

El primer grupo que se establece es referente a la variable edad; estableciendo rangos: el primero es para menores de edad, -que va de los 12 a los 17 años-, donde la media de ingresos es de \$2523.81 mensuales, resultando el rango de edad que registra ingresos más bajos; el segundo rango es para los trabajadores que tienen entre 18 y 45 años, cuyo ingreso promedio es de \$ 6179.13 -valor cercano a la media general-; en el tercer rango se encuentran los que van de 46 a 65 años, donde podemos darnos cuenta que la media es la más alta del grupo, la cual es de \$ 7081.32; finalmente, en el cuarto rango, se encuentran los trabajadores de entre 66 y 98 años, cuyo promedio de ingresos es de \$ 5438.10.

Tabla 2. Ingresos mensuales promedio considerando características del trabajador

Subgrupos	Promedio ingreso mensual
Total	\$6,275.42
Edad (años)	
12-17	\$2,523.81
18-45	\$6,179.14
46-65	\$7,081.32
66-98	\$5,438.10
Sexo	
Hombre (1)	\$6,295.03
Mujer (0)	\$6,247.98
Años de escolaridad	
0	\$5,907.41
1-6	\$6,281.21
7-9	\$6,350.15
10-12	\$6,231.64
13-17	\$6,231.64
18-24	\$6,507.80
Estado conyugal	
Casado/ unión libre	\$6,257.67
Cualquier otro	\$6,301.36

Fuente: Elaboración propia con datos de la ENOE, INEGI.

Dentro de la clasificación que se realiza con las características del trabajador, la edad es la variable que genera principalmente diferencias salariales; las categorizaciones por sexo, años de escolaridad y estado conyugal no conducen a desigualdades significativas. Cabe resaltar que de acuerdo a los preceptos de la teoría del capital humano, se esperaría que al catalogar por años de escolaridad, los ingresos de las personas aumenten conforme halla más preparación académica; sin embargo, la evidencia empírica no es acorde a lo esperado por dicha teoría.

Tabla 3. Ingresos mensuales promedio considerando características del empleo

Subgrupos	Promedio ingreso mensual
Sindicato	
Sí (1)	\$8,989.00
No (0)	\$5,734.92
Contrato	
Sí (1)	\$7,694.66
No (0)	\$4,287.90
Tamaño de la empresa	
Micro	\$4,356.74
Pequeña	\$6,731.56
Mediana	\$7,234.61
Grande	\$8,678.68
Rama	
Agropecuario	\$6,350.53
Industria manufacturera	\$6,221.03
Construcción	\$6,350.53
Comercio	\$6,252.80
Servicios	\$6,306.75

Fuente: Elaboración propia con datos de la ENOE, INEGI.

Posteriormente, en el análisis por características del empleo (Tabla 3), se prescinde de las variables vacaciones, utilidades, aguinaldo, crédito a la vivienda y seguridad social, debido a que los resultados son semejantes a la variable contrato. Los trabajadores sindicalizados ganan en promedio 36% más que aquellos que no están adscritos a un sindicato. Las personas que en su empleo cuentan con un contrato escrito, además de tener más estabilidad laboral alcanzan ingresos mayores en un 45% respecto a los que no tienen contrato.

En promedio, las empresas grandes pagan a sus empleados casi el doble de lo que pagan las empresas micro, es decir, el tamaño de la empresa es una variable que produce diferenciación salarial, contrario a lo que sucede cuando se analiza por rama de actividad económica, ya que no se encuentran variaciones relevantes.

Tabla 4. Ingresos mensuales promedio considerando la ciudad de residencia del trabajador

Subgrupos	Promedio ingreso mensual
Ciudad¹	
La Paz	\$8,017.84
Villahermosa	\$7,850.27
Hermosillo	\$7,795.79
Monterrey	\$7,257.91
Saltillo	\$7,120.34
Guadalajara	\$6,347.33
Tijuana	\$6,139.83
México	\$6,075.47
Toluca	\$6,014.52
Durango	\$5,587.48
León	\$5,519.18
Cuernavaca	\$5,293.70
Acapulco	\$4,978.88
Tlaxcala	\$4,514.00
Tamaño de localidad	
Grande (1)	\$6,267.86
Otra (0)	\$6,328.76

Fuente: Elaboración propia con datos de la ENOE, INEGI.

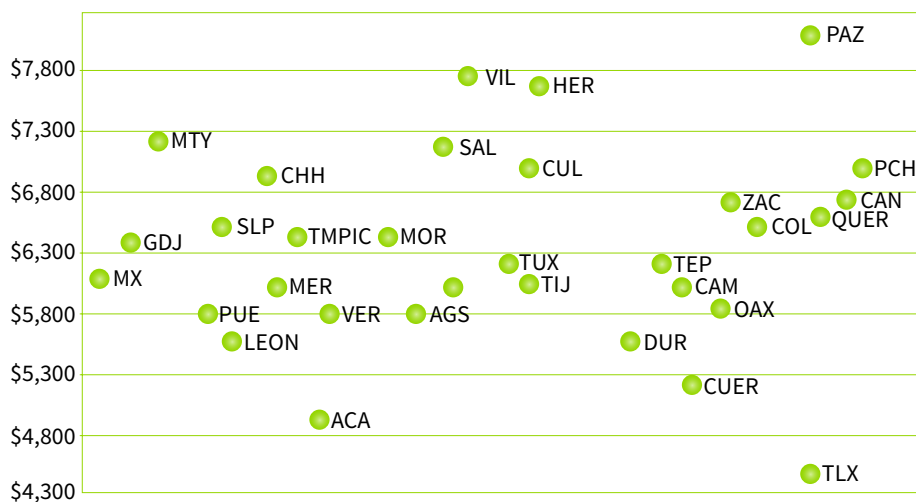
Al generar un subgrupo por la ciudad en la que reside el trabajador, los ingresos mensuales son muy diferentes (Tabla 4); por ejemplo, un trabajador que vive en la ciudad de La Paz gana en promedio \$8,017.84, mientras que un empleado que resida en la ciudad de Tlaxcala alcanza ingresos de \$4,514.00 mensuales. Este hecho muestra que la localización es un factor determinante al momento de establecer niveles salariales. El vivir en una localidad grande no aporta incrementos salariales importantes.

En la Gráfica 3 se refleja el hecho de que existen grandes desigualdades en los ingresos al ser considerados por ciudades². En promedio, los trabajadores de las ciudades de La Paz, Hermosillo y Villahermosa ganan más que los trabajadores del resto de las ciudades de México.

¹ Se seleccionaron sólo algunas de las 32 ciudades autorepresentadas que considera la ENOE.

² Para mayor detalle de la lista de ciudades y nombres completos, consultar el Anexo 1.

Gráfica 3. Ingresos mensuales promedio por ciudad



Fuente: Elaboración propia con datos de la ENOE, INEGI.

Modelo econométrico multinivel

La técnica econométrica multinivel ha recobrado mucha importancia en los últimos años, debido a las aplicaciones que se tienen en el análisis de datos anidados. La idea de la estructura jerárquica es que las unidades aparecen agrupadas y que varios grupos forman grupos en otro nivel y así sucesivamente; por eso se les llama también modelos lineales multinivel (Goldstein, 1987, p.70).

Los sujetos que pertenecen al mismo grupo reciben una serie de influencias comunes, por lo tanto, las observaciones dentro de un mismo grupo no son independientes entre sí, y si se aplican los métodos de análisis de regresión tradicionales, se viola uno de los principios básicos de econometría: la independencia de los datos, lo que se traduce en la subestimación de los errores estándar de los coeficientes de regresión.

El análisis multinivel permite incluir características de grupos diferentes en modelos de comportamiento individual. Al incorporar las características de grupos al modelo multinivel, se tiene en cuenta la estructura jerárquica de los datos, y por lo tanto, se obtienen estimaciones correctas de los errores estándar.

En datos estructurados jerárquicamente, los individuos pertenecientes al mismo grupo son probablemente más parecidos entre sí que individuos en grupos diferentes. Debido a esto, las variaciones en el resultado se pueden deber a diferencias entre grupos y a diferencias individuales dentro del mismo grupo.

Una jerarquía consiste en un conjunto de observaciones que conforman el primer nivel o micro nivel que se encuentran anidadas dentro de un nivel superior, denominado segundo nivel o nivel macro.

Las variables en estos modelos se diferencian por la unidad de análisis en la que se miden dichas variables. A las que se extraen de la unidad más pequeña se les denomina variables a nivel individual; aquellas que miden a las unidades que agrupan a otras unidades, por ejemplo, las áreas geográficas, se les llama variables a nivel grupo o variables contextuales. En el caso general, las observaciones individuales i , están anidadas en unidades agregadas j .

La estructura que se considera para el modelo multinivel es la siguiente (ver Figura 1):

Nivel 1

Variable dependiente: Ingresos mensuales (es el ingreso de cada individuo i , es decir, de cada trabajador).

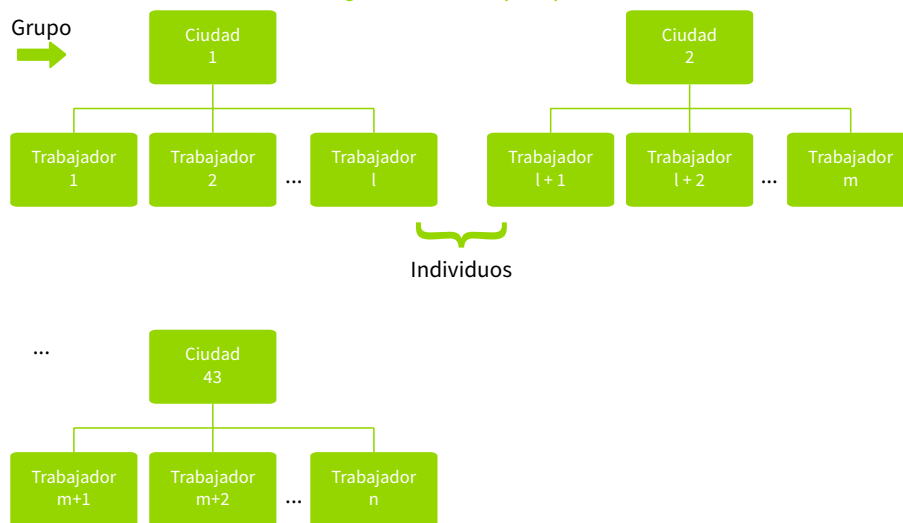
Variables explicativas: Edad, sexo, años de escolaridad, estado conyugal, sindicato, contrato, aguinaldo, vacaciones, utilidades, crédito a la vivienda, tamaño de la empresa, rama, seguridad social y tamaño de la localidad.

Nivel 2

Ciudad (a nivel grupo, es decir, se refiere a la ciudad en la que vive el trabajador encuestado).

Debido a que se pretende identificar los factores que generan diferenciación salarial, se consideran los ingresos mensuales de los asalariados como la variable dependiente. La estructura jerárquica que se propone es la de trabajadores anidados en ciudades, con características propias de cada individuo como: sexo, años de escolaridad, etc., y características propias del trabajo, tales como: tamaño de la empresa, si cuenta con seguridad social, etc. Se asume que las características del Nivel 2, es decir, la ciudad en la que reside el trabajador, mediará los efectos sobre el nivel de ingreso. Dicha estructura significa que la localización del individuo genera efectos diferenciales al momento de fijar los salarios.

Figura 1. Estructura jerárquica



Fuente: Elaboración propia

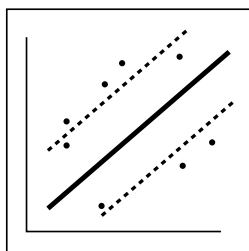
El coeficiente de correlación intraclase es un indicador de la homogeneidad interna de los grupos. Mide la similitud de las unidades del nivel individual y de las diferencias entre las unidades del nivel macro. Siguiendo a Cebolla (2013, pp. 87-88) el valor de rho (ρ), también llamado correlación intraclase (ICC por sus siglas en inglés *Interclass Correlation Coefficient*, o VPC *Variance Partition Coefficients*), debe ser interpretado como un coeficiente de correlación intraunidad agregada, una medida del grado de homogeneidad que describe a las unidades individuales agrupadas en una misma unidad de anidamiento. De forma asimétrica, el coeficiente de correlación intraclase es $1 - \rho$.

En la Figura 2, cuando ρ es grande (cercano a 1, como en el caso A), una mayor parte de la varianza de la variable dependiente se debe a factores relacionados con los grupos. Esto quiere decir que dentro de cada grupo j (descrito por las rectas discontinuas) las unidades individuales se parecen mucho entre sí. En caso de que suceda lo contrario y ρ sea cercano a cero (como en el escenario C), las unidades individuales dentro de cada grupo j son muy heterogéneas. Un valor pequeño de ρ sugiere que podemos ignorar la estructura multinivel, ya que, en sentido estricto, la proporción de la varianza a explicar en el nivel de los grupos es muy escasa. Aquí la mayor parte de la explicación parece deberse a los individuos. Aunque esta apreciación sea conceptualmente

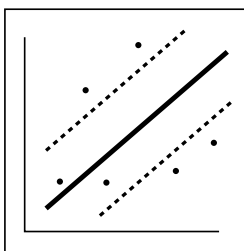
correcta, la opción más ortodoxa no es decidir sobre la pertinencia de modelos multinivel a partir de la información que nos proporciona p ; es necesario hacer un contraste que contraponga el ajuste de dos modelos, uno que considere dos niveles de varianza y otro que no (Cebolla, 2013, pp. 87-89).

Figura 2. Interpretación del coeficiente de correlación entre clases

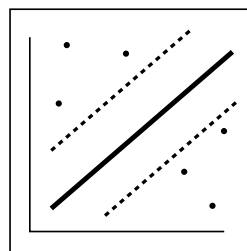
A. Dependencia máxima de los grupos ($p=1$)



B. Dependencia media de los grupos ($p=0.5$)



C. Mínima dependencia de los grupos ($p=0$)



Fuente: Cebolla (2013)

Si se establece una jerarquía de dos niveles, donde los individuos del nivel 1 se encuentran agrupados dentro de unidades del nivel 2, se incorpora el subíndice j en el modelo, y adicionalmente, se incorporan variables independientes en el nivel 1. La ecuación de regresión para el i -ésimo individuo dentro del j -ésimo grupo (elementos de segundo nivel), está dada por:

$$Y_{ij} = \beta_{0j} + \beta_{1j} x_{ij} + e_{ij}$$

$$\beta_{0j} = \gamma_{00} + U_{0j}$$

$$\beta_{1j} = \gamma_{10} + U_{1j}$$

En una ecuación:

$$Y_{ij} = \gamma_{00} + \gamma_{10} x_{ij} + U_{0j} + U_{1j} + e_{ij}$$

En la regresión de un nivel se tiene la variabilidad asociada sólo al término residual; en modelos de dos niveles se presenta la variabilidad entre los individuos y entre los grupos; por lo que se tiene la posibilidad de modelar la variabilidad no explicada de:

Nivel 1 $\rightarrow e_{ij}$ la variabilidad dentro grupos.
 Nivel 2 $\rightarrow U_{0j}, U_{1j}$ la variabilidad entre grupos.

La variabilidad en el segundo nivel es la característica principal de los modelos multinivel. El objetivo de la modelización multinivel es considerar dentro de un mismo modelo los distintos niveles de la estructura jerárquica y conocer la variabilidad debida al segundo nivel.

Coficiente de correlación intraclase residual:

$$\text{var}(Y_{ij} | x_{ij}) = \text{var}(U_{0j} + e_{ij}) = \text{var}(U_{0j}) + \text{var}(e_{ij}) = \tau_0^2 + \sigma^2$$

$$\text{cov}(Y_{ij}, Y_{i'j} | x_{ij}, x_{i'j}) = \text{var}(U_{0j}) = \tau_0^2$$

$$\rho(Y_{ij}, Y_{i'j} | x_{ij}, x_{i'j}) = \frac{\tau_0^2}{\tau_0^2 + \sigma^2}$$

Los estimadores más utilizados en el análisis de regresión multinivel son los estimadores de Máxima Verosimilitud (*Maximun Likelihood Estimators*, MLE). El objetivo de la estimación de este método es encontrar un estimador del parámetro, que sea lo más cercano al verdadero valor del parámetro (Hox, 2002, pp. 257-259); es decir, que de un conjunto de datos la estimación de Máxima Verosimilitud toma el valor del parámetro que da lugar a la distribución con la que los datos son más probables.

El procedimiento de máxima verosimilitud produce errores estándar para la mayoría de las estimaciones; la significancia de un predictor viene dada por la razón entre el estimador del parámetro y su error típico. La prueba es el Test de Wald, que sirve para aceptar o rechazar la hipótesis nula establecida sobre el estimador del parámetro β :

$$H_0 = \beta = 0$$

Se verifica que:

$$\frac{\hat{\beta}}{\hat{s}_\beta} \sim N(0,1) \text{ o lo que es equivalente } \left(\frac{\hat{\beta}}{\hat{s}_\beta} \right)^2 \sim X_1^2 \text{ g. l.}$$

En la práctica, una regla general para determinar la significancia de un predictor es que si,

$$z = \frac{\text{parámetro}}{\text{error estándar}} > 2 \rightarrow p < 0.05$$

El análisis del estadístico *Deviance* es una generalización del análisis de la varianza de una serie de modelos anidados -cada uno incluyendo más términos que los anteriores-. Dada una sucesión de modelos anidados, este estadístico se usa como una medida de discrepancia entre el modelo original y el que incorpora más elementos.

Esta prueba contrasta la ganancia que representa considerar en el modelo la estructura jerárquica de los datos con respecto a un modelo especificado de forma idéntica que solo considera la varianza en un único nivel. La prueba asociada se hace calculando la diferencia entre los logaritmos de la verosimilitud (*log likelihood values*) de los modelos multiplicada por dos.

$$LR = 2 (\log \text{verosimilitud modelo 2} - \log \text{verosimilitud modelo 1})$$

Los procedimientos de Máxima Verosimilitud producen el estadístico *Deviance* (Hox, 2002, p.43), el cual indica qué tan bien se ajusta el modelo a los datos:

$$Dev = -2 \ln (L_1) = -2 \log (\text{likelihood})$$

En donde L_1 , es el valor del máximo de la función de verosimilitud (*likelihood*) en la estimación de los parámetros del modelo 1.

Si en el modelo nulo no hay varianza estadísticamente distinta de cero en los niveles contemplados, ningún modelo que se derive de éste añadiendo variables explicativas mejorará el ajuste, ya que la varianza del intercepto no es significativamente distinta de cero. Si dos modelos están anidados, es decir, un modelo se obtiene a partir de otro más general eliminando parámetros de este modelo, entonces podemos compararlos, para lo cual se utiliza el estadístico *Deviance* de cada modelo, pues la diferencia se distribuye como una chi-cuadrada con grados de libertad iguales a la diferencia del número de parámetros estimados en los modelos que se están comparando bajo la hipótesis nula:

$$H_0 = \text{ambos modelos son iguales}$$

La significación estadística de esta prueba es una confirmación de que la estimación multinivel es la más apropiada; esto es, si se rechaza la hipótesis nula, la regresión multinivel modela mejor el fenómeno en cuestión.

Resultados de la estimación

El primer modelo incluye todas las variables descritas en la Tabla 1. No todas las variables son estadísticamente diferentes de cero, ya que al hacer pruebas de significancia individual se concluye que se tiene que prescindir de las variables: sexo, años de escolaridad, estado conyugal, aguinaldo, seguridad social, las cinco ramas económicas y el tamaño de la localidad (ver Figura 3).

Figura 3. Modelo de dos niveles (primera estimación)

Mixed-effects ML regression	Number of obs	=	23875		
Group variable: ciudad	Number of groups	=	32		
	Obs per group: min	=	389		
		avg	= 746.1		
		max	= 1153		
Log likelihood = -234973.1	Wald chi2 (18)	=	4804.77		
	Prob > chi2	=	0.0000		
Ing_mensual	Coef.	Std. Err.	z	P>[z]	[95% Conf. Interval]
edad	45.00444	2.376185	18.94	0.000	40.34721 49.66168
sexo	59.40577	62.62052	0.95	0.343	-63.3282 182.1397
anos_esc	5.454286	7.010562	0.78	0.437	-8.286162 19.19474
edo_conyugal	-39.35904	61.04286	-0.64	0.519	-159.0009 80.28277
sindicato	981.9343	89.47371	10.97	0.000	806.5691 1157.3
contrato	757.372	106.7812	7.09	0.000	548.0847 966.6592
aguinaldo	152.0754	111.3396	1.37	0.172	-66.14626 370.297
vacaciones	980.8161	123.8843	7.92	0.000	738.0073 1223.625
utilidades	-679.3629	85.72151	-7.93	0.000	-847.374 -511.3518
cred_viv	483.1127	112.2339	4.30	0.000	263.1382 703.0871
tamano_emp	248.1549	11.88236	20.88	0.000	224.8659 271.4439
seg_social	141.7818	130.4819	1.09	0.277	-113.9579 397.5216
agrop	203.2426	439.9879	0.46	0.644	-659.1179 1065.603
ind_manuf	-123.0299	315.0139	-0.39	0.696	-740.4457 494.3859
construc	-70.74122	324.2739	-0.22	0.827	-706.3063 564.8239
comercio	-88.67883	312.9409	-0.28	0.777	-702.0317 524.674
servicios	15.42391	308.7126	0.05	0.960	-589.6418 620.4896
tam_loc	-24.27072	96.66419	-0.25	0.802	-213.729 165.1876
_cons	1678.053	361.5096	4.64	0.000	969.5072 2386.599

Random_effects Parameters	Estimate	Std. Err.	[95% Conf. Interval]
ciudad: Identity			
var(_cons)	292584.3	80705.99	170395 502394.7
var(Residual)	2.06e+07	188988	2.03e+07 2.10e+07
LR test vs. linear regression: $\chi^2_{(01)} = 237.83$ Prob >= $\chi^2_{(01)} = 0.0000$			
. xtmrho			
Levels: ciudad			
level 1			
Intraclass correlation (ICC): $\rho_{01} = 0.01398$			

Fuente: Elaboración propia, con datos de la ENOE, INEGI.

Además se estimó un modelo de un solo nivel -regresión calculada por MCO-, modelo de dos niveles vacíos, y modelos de dos niveles incorporando de una en una variables independientes en el primer nivel. En cada uno de los casos se construyó una prueba de hipótesis y se aplicó el criterio del estadístico *Deviance* para seleccionar las variables explicativas. Dando como resultado el modelo de la Figura 3.

En modelo final todas las variables son significativas y tienen un efecto positivo sobre la variable dependiente, excepto las utilidades que arrojan un coeficiente negativo (ver Figura 4). De las características propias del trabajador consideradas en esta investigación, la edad es la única variable que contribuye a explicar el comportamiento de los ingresos de los trabajadores. El que una empresa ofrezca a sus empleados la opción de pertenecer a un sindicato, que le otorgue un contrato laboral escrito, vacaciones con goce de sueldo, acceso a crédito a la vivienda y que la empresa sea de tamaño grande, se refleja en ingresos mayores; es decir, el pertenecer a alguno de los subgrupos mencionados aporta diferencias salariales positivas respecto a aquellos individuos que no están dentro de estos subgrupos.

Figura 4. Modelo de dos niveles (estimación final)

Mixed-effects ML regression			Number of obs =		23875	
Group variable: ciudad			Number of groups =		32	
			Obs per group: min =		389	
			avg =		746.1	
			max =		1153	
Log likelihood = -234978.06			Wald chi2 (7) =		4792.62	
			Prob > chi2 =		0.0000	
Ing_mensual	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
edad	45.1684	2.373994	19.03	0.000	40.51546	49.82135
sindicato	982.1084	89.46273	10.98	0.000	806.7646	1157.452
contrato	807.9458	101.6508	7.95	0.000	608.7139	1007.178
vacaciones	1104.695	103.1282	10.71	0.000	902.5675	1306.823
utilidades	-681.5819	85.70742	-7.95	0.000	-849.5653	-513.5984
cred_viv	554.9758	99.04568	5.60	0.000	360.8498	749.1017
tamano_emp	250.1951	11.7777	21.24	0.000	227.1112	273.2789
_cons	1723.005	140.5722	12.26	0.000	1447.489	1998.522

(Continúa en la siguiente página) ►

Random_effects Parameters	Estimate	Std. Err.	[95% Conf. Interval]
ciudad: Identity			
var(_cons)	296509.8	81660.49	172827.7 508703.6
var(Residual)	2.06e+07	189063.3	2.03e+07 2.10e+07
LR test vs. linear regression: $\chi^2(01) = 242.50$ Prob >= $\chi^2 = 0.0000$			
. xtmrho			
Levels: ciudad			
level 1			
Intraclass correlation (ICC): $\rho_{01} = 0.01416$			

Fuente: Elaboración propia, con datos de la ENOE, INEGI.

En promedio hay 746 individuos por ciudad, en el grupo más pequeño existen 389 trabajadores, mientras que en el grupo más grande están registrados 1153.

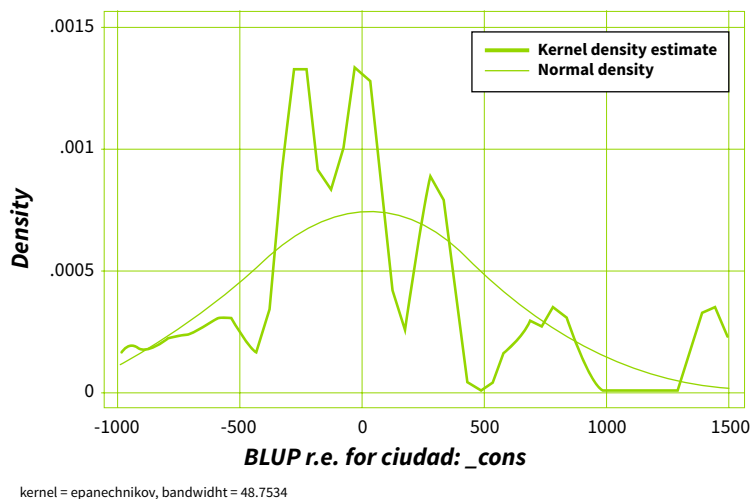
El coeficiente de correlación intraclase (ICC) tiene un valor igual a 0.014, lo cual significa que el 1.4% de la variación ocurre en el segundo nivel, debido a que la fórmula del ICC estima la proporción de varianza que se debe al nivel agregado.

La constante adquiere un valor de $\gamma_{00}=1,723.005$, que representa el promedio de ingresos a través de todas las ciudades, es la gran media o media anidada (\$1,723); por ejemplo, si un individuo pertenece a un sindicato la media de sus ingresos aumenta en \$982.10.

Dado que la ciudad en la que reside el individuo ha sido tomada como la variable del nivel agregado (nivel 2), se dice que el pertenecer a una ciudad en específico aporta diferencias en los salarios; es decir, los trabajadores que viven en ciudades como La Paz, obtienen ingresos superiores a otros individuos con características semejantes pero que habitan en otra ciudad. Las características de la ciudad mediarán los efectos sobre el nivel de ingreso.

Se genera una nueva variable denominada U_{0j} , que registra para cada grupo la perturbación aleatoria. Los efectos aleatorios representan información muy útil para describir las diferencias entre los grupos, ya que cuantifican la distancia entre la media general (γ_{00}) de los ingresos de las ciudades y el ingreso de la ciudad j . La Figura 5 permite comprobar cuánto se desvía de la normalidad la distribución de este parámetro (Cebolla, 2013, p. 89).

Figura 5. Densidad Kernel



Fuente: Elaboración propia, con datos de la ENOE, INEGI.


Conclusiones

En el presente artículo se muestra evidencia de que existen diferencias salariales en el mercado de trabajo mexicano. El análisis estadístico, gráfico y econométrico da como resultado que la determinación de los ingresos de los trabajadores es dada por características consideradas en tres rubros:

- a) Características propias del trabajador: la edad. Se concluye que el pertenecer a un grupo de edad mayor da como resultado un ingreso superior.
- b) Características del trabajo: sindicato, contrato, vacaciones, utilidades, crédito a la vivienda y el tamaño de la empresa. Pertenecer a cualquiera de estos grupos aporta un efecto positivo al momento de determinar el salario; excepto con las utilidades, las cuales presentan un coeficiente negativo en el modelo.
- c) Características del lugar de residencia del trabajador: ciudad. El que un asalariado viva en una ciudad en particular le aporta un efecto dife-

rencial en su ingreso. Los trabajadores están anidados en ciudades, de tal forma que el pertenecer a uno de los grupos -ciudades-, aporta un efecto positivo o negativo en el salario respecto al promedio total de los ingresos de los trabajadores. La localización, por lo tanto, contribuye al nivel salarial.

La teoría del capital humano sostiene que la educación es la diferencia determinante en los salarios de individuos con características semejantes; sin embargo, los datos no muestran evidencia para sostener dicha afirmación, debido que al clasificar los años de escolaridad en rangos, no se encontraron variaciones importantes en los ingresos; además, esta variable no resultó significativa en los modelos estimados. Dicha teoría espera una correlación positiva y significativa entre los años de escolaridad y el ingreso de los trabajadores de los individuos; los resultados de esta investigación no confirman tal relación.

Resulta relevante atender políticas económicas que coadyuven a frenar las brechas salariales de los trabajadores, atendiendo las variables que resultaron significativas en la presente investigación. 

Anexo

Anexo 1. Lista de ciudades

Clave ³	Ciudad	Abreviatura
01	México	MX
02	Guadalajara	GDJ
03	Monterrey	MTY
04	Puebla	PUE
05	León	LEON
07	San Luis Potosí	SLP
08	Mérida	MER
09	Chihuahua	CHH
10	Tampico	TMPIC
12	Veracruz	VER
13	Acapulco	ACA
14	Aguascalientes	AGS

(Continúa en la siguiente página) ►

³ De acuerdo a la ENOE.

Clave ³	Ciudad	Abreviatura
15	Morelia	MOR
16	Toluca	TOL
17	Saltillo	SAL
18	Villahermosa	VIL
19	Tuxtla Gutiérrez	TUX
21	Tijuana	TIJ
24	Culiacán Rosales	CUL
25	Hermosillo	HER
26	Durango	DUR
27	Tepic	TEP
28	Campeche	CAM
29	Cuernavaca	CUER
31	Oaxaca	OAX
32	Zacatecas	ZAC
33	Colima	COL
36	Querétaro	QUER
39	Tlaxcala	TLX
40	La Paz	PAZ
41	Cancún	CAN
43	Pachuca	PCH

Fuente: Elaboración propia, con información de la ENOE, INEGI

Referencias

- 🍃 CEBOLLA, BOADO HÉCTOR (2013). *Introducción al análisis multinivel*. España: Editorial Centro de Investigaciones Sociológicas, Vol. 49 de Cuadernos Metodológicos.
- 🍃 CUE, AGUSTÍN Y LUIS QUINTANA (2009). *Introducción a la Microeconomía: Un enfoque integral para México*. México: Grupo Editorial Patria.
- 🍃 Diario Oficial de la Federación (2015, junio). Ley Federal del trabajo. Recuperado el 1 de octubre de 2015, de http://www.diputados.gob.mx/LeyesBiblio/pdf/125_120615.pdf.
- 🍃 GARZA, ACEVEDO BLANCA (2014). *Habilidades, salarios y desigualdad regional en México*: UNAM, FES Acatlán. Tesis de doctorado en Economía.
- 🍃 GOLDSTEIN, HARVEY (1987). *Multilevel models in educational and social research*. Londres: Charles Griffin & Company.

- HOX, JOOP J. (2002). *Multilevel analysis: Techniques and applications*. (2a ed.). Estados Unidos: Lawrence Erlbaum Associates, Inc.
- Instituto Nacional de Estadística y Geografía, Encuesta Nacional de Ocupación y Empleo (ENOE). Recuperado el 7 de septiembre de 2015, de <http://www.inegi.org.mx/est/contenidos/proyectos/encuestas/hogares/regulares/enoe/default.aspx>.
- SCHULTZ, TEODORE (1961). Investment in Human Capital. *American Economic Review*, Vol. 51, Núm. 1, 1-17.
- (1985, diciembre). *Investing in people: the economics of population quality*. España: Ariel, S.A.

JÉSSICA GÁMEZ ARROYO es maestra y licenciada en Economía por la UNAM FES Acatlán, Profesora de asignatura en la FES Acatlán de 2010 a 2014 y profesora de asignatura en la UAEM Valle de México. Ha participado en varios congresos y seminarios nacionales. Temas de especialización: diferenciación salarial, modelos econométricos, econometría regional.

EDUARDO ROSAS ROJAS es Actuario por la UAEM, Especialista en Finanzas Públicas, Maestro y Doctor en Economía por la UNAM. Profesor investigador, categoría C y Coordinador de la Licenciatura en Actuaría del Centro Universitario UAEM Valle de México, también participa como Profesor del Posgrado en Economía de la UNAM. Miembro del Cuerpo Académico UAEM-CA.96, con registro ante el PRODEP (SEP). Ha publicado artículos en revistas arbitradas e indexadas, así como capítulos de libro. Líneas de investigación: modelos de crecimiento económico, y modelado de políticas fiscales y cambiarias.