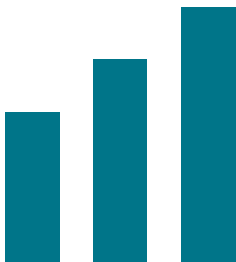


DIFERENCIAS DEL INGRESO ENTRE TRABAJADORES DE LOS SECTORES FORMAL E INFORMAL DE LA ECONOMÍA GUATEMALTECA EN EL AÑO 2018

Síntesis de tesis



Lcdo. Joaquín Isaí Alvarado López

Licenciado en Economía Empresarial por la Universidad Rafael Landívar

Correo: joaquinisai3506@yahoo.es

Fecha de recepción: 2 de julio de 2020

Fecha de aceptación: 25 de agosto de 2020

Resumen

El presente artículo pretende determinar las diferencias significativas entre el ingreso de los trabajadores formales e informales, mediante tres modelos econométricos. El primer modelo es el método de Heckit, que corrige el sesgo de selección; el segundo modelo es la ecuación mincenaria de ingresos de Mincer, que estima el impacto de un año adicional de estudio en los ingresos laborales de los individuos; y por último, el modelo de descomposición Oaxaca-Blinder, que estima dos regresiones lineales del salario en función a las características observables midiendo las diferencias y la brecha salarial. La base de datos utilizada para

las estimaciones se obtuvo de la Encuesta Nacional de Empleo e Ingresos del 2018. Se concluyó que existen diferencias significativas en el ingreso de los trabajadores formales e informales, y la brecha salarial explicada por la descomposición Oaxaca-Blinder es de un 99.89 %. El 32.63 % es explicado por las diferencias en las dotaciones, el 72.05 % se atribuye a las diferencias en los coeficientes y el 0.05 % restante se explica por la interacción entre los coeficientes y las dotaciones.

Palabras clave: trabajadores formales, trabajadores informales, salario, ingresos, brecha salarial, dotaciones

INCOME DIFFERENCES BETWEEN WORKERS OF THE FORMAL AND INFORMAL SECTOR OF GUATEMALA'S ECONOMY DURING 2018

Scientific research paper

Abstract

The present article pretends to determine significant income differences between formal and informal workers, using three different econometric models: the Heckit model, which corrects biased selections; the Mincer equation model, in which we calculate the impact that one additional year of studies has on wage income; and finally the Oaxaca-Blinder decomposition model, which estimates two salary linear regressions in terms of observable characteristics by measuring differences and salary gaps. The database used for these calculations draws on the 2018 National Employment and Income Survey. The study concludes that there are important income differences between formal and informal workers. The Oaxaca-Binder decomposition model portrays a 99.89% wage gap. Of this gap, 32.63 % is attributed to allocation differences, 72.05% to coefficient differences and the remaining 0.05% is attributed to the interaction of coefficients and allocations.

Keywords: *formal workers, informal workers, wages, income, wage gap, endowments*

1. Introducción

La presente investigación aborda el tema de las diferencias que existen en el ingreso laboral de los trabajadores que forman parte de los dos grandes sectores de la economía guatemalteca, el sector informal y el sector formal, para el 2018. Esto se hizo por medio de los datos proporcionados por el Instituto Nacional de Estadística (INE), a través de la Encuesta Nacional de Empleo e Ingresos (ENEI) presentada en el 2018.

El objetivo de la investigación fue conocer las circunstancias actuales en las que se encuentra el mercado laboral guatemalteco, como también la conexión con otros factores que la teoría económica señala como determinantes sobre la desigualdad salarial que existe en el país. Como complemento, se buscó lograr establecer si estas relaciones son un factor que determina la existencia de informalidad o formalidad en Guatemala.

Para estudiar las diferencias de dichos sectores, se hace uso de un modelo econométrico basado en la metodología de descomposición citada por los autores Blinder (1973) y Jann (2008), quienes estiman dos regresiones lineales del salario en función a las características observables como: género, edad, nivel de escolaridad, experiencia, entre otras; una regresión para cada uno de los dos sectores.

2. Modelos y datos

Para el presente estudio, los sujetos de investigación están conformados por los estratos de población ocupada que representan el sector formal y el sector informal de la economía guatemalteca. En la ENEI 2018 se incluye una variable que especifica si los individuos pertenecen al sector formal o al informal; por consiguiente, la población ocupada del país se encuentra en 7 003 398 personas. Según estos datos, el sector informal está formado por el 69.5 % de la población ocupada, que representa 4 867 361 personas, mientras que el sector formal está formado por el restante de 30.5 %, que representa 2 136 036 personas.

Sin embargo, las estimaciones pueden verse afectadas por sesgo de selección, esto significa que los individuos que se reportan en las encuestas se autoseleccionaron para entrar en el sector formal o informal. Si las características de estos difieren sistemáticamente, el modelo haría estimaciones de una muestra no aleatoria y los resultados no serían representativos para toda la población; por lo que se implementará el modelo Heckit (Perodi, Ramírez y Thompson, 2017).

Para definir la población, se utilizaron los datos proporcionados por el INE. De acuerdo con las proyecciones y estimaciones de población para el 2018, se estima que la población de Guatemala es de 17 311 086 habitantes.

El esquema de muestreo que adopta la ENEI 2018 es de carácter probabilístico, bietápico, estratificado y por conglomerados. El tamaño de la muestra fue calculado considerando como variable principal la tasa de subempleo visible (TSV). Este tamaño de la muestra es válido para dar estimaciones a nivel de cada dominio de estudio y total del país.

La muestra en su total fue de 5277 viviendas en 363 sectores: de la misma, 40 sectores pertenecen al dominio urbano metropolitano, 33 al resto urbano, 11 al rural nacional y 25 al nacional, con un nivel de confiabilidad del 95 %. La submuestra fue de 8573 observaciones, de las cuales 3230 pertenecen al sector formal y 5343 pertenecen al sector informal.

La presente investigación es de tipo descriptivo, y para calcular las diferencias salariales entre los trabajadores del sector formal e informal, se utilizarán tres modelos econométricos.

El primer modelo es el método de Heckman (1979) o modelo Heckit, el cual corrige el sesgo de selección en dos etapas; este error puede atribuirse a la omisión de una variable relevante en el modelo. En la primera etapa se estima la variable omitida, conocida como la razón inversa de Mills, por medio de un modelo *probit*

que estima la probabilidad de que un individuo participe en el sector formal o en el sector informal.

En la segunda etapa, se incluye la razón inversa de Mills al modelo de regresión de Mínimos Cuadrados Ordinarios (MCO) como variable explicativa en el modelo (Perodi, Ramírez y Thompson, 2017). Para el cálculo de la razón inversa de Mills se estimó la ecuación de selección de la siguiente manera:

$$P = \begin{cases} 1 & \text{si } P_i^* = \beta x_i^* + \omega_i \\ 0 & \text{si lo contrario} \end{cases}$$

En donde el error estándar ω_i tiene una distribución normal, y x_i^* es un vector de características observables de los individuos que determinan su propensión a trabajar en el sector formal o en el sector informal. A partir de esta ecuación se arma la razón inversa de Mills (λ), el selector de selectividad λ_i que se forma a través del cociente de la función de densidad de probabilidad θ_i y la función de probabilidad acumulada Θ_i (Rizzo, 2004). Entonces:

$$\lambda_i = \frac{\theta_i}{\Theta_i}$$

En la segunda etapa, se estima una ecuación por MCO incluyendo como una variable más el regresor de selectividad, la razón inversa de Mills (λ) (Rizzo, 2004). La ecuación del modelo es la siguiente:

$$W_i = \beta_1 + \beta_2 X + \lambda_i + \mu$$

Donde:

W_i = logaritmo natural del ingreso laboral de los individuos

β_1 = coeficiente, intercepción con el eje Y

$\beta_2 X$ = características observables de los individuos

λ_i = la razón inversa de Mills

μ = errores

Si (λ) resulta estadísticamente significativa, la muestra es considerada no aleatoria y las estimaciones estarían sesgadas de no incluir (λ) como variable explicativa en el modelo.

El segundo modelo es la ecuación mincenaria de ingresos de Mincer (1974), para estimar el impacto de un año adicional de estudio en los ingresos laborales de los individuos.

La ecuación de Mincer estima por medio de MCO un modelo de tipo semilogarítmico, en el cual se utiliza como variable dependiente el logaritmo de los ingresos salariales para el sector formal e informal, y como variables explicativas los años de escolaridad, la experiencia laboral y la misma al cuadrado (Álvarez y Teijeiro, 2008).

$$W_F = \beta_{F1} + \beta_{F2} S + \beta_{F3} EXP + \beta_{F4} EXP^2$$

$$W_{inf} = \beta_{inf1} + \beta_{inf2} S + \beta_{inf3} EXP + \beta_{inf4} EXP^2$$

Donde:

W_F = logaritmo natural del ingreso laboral de los trabajadores formales

W_{inf} = logaritmo natural del ingreso laboral de los trabajadores informales

β_{inf1} = coeficiente, intercepción con el eje Y

$\beta_{inf2} S$ = número de años de educación formal completada para los trabajadores informales

β_{F1} = coeficiente, intercepción con el eje Y

$\beta_{F2} S$ = número de años de educación formal completada para los trabajadores formales

$\beta_{F3} EXP$ = años de experiencia laboral para los trabajadores formales

$\beta_{inf3} EXP$ = años de experiencia laboral para los trabajadores informales

$\beta_{inf3} EXP$ = años de experiencia laboral para los trabajadores informales al cuadrado

$\beta_{F4} EXP^2$ = años de experiencia laboral para los trabajadores formales al cuadrado

Este modelo tiene una estructura muy simple que no caracteriza de forma completa la distribución del ingreso, edad y escolaridad, pero se sigue utilizando como un punto de partida para construir modelos de mayor complejidad como lo es en este análisis (Metzger, 2006). El modelo mincenario es la idea fundamental para el modelo final, debido a que aportará distintas variables que podrán ser utilizadas como características observables en el segundo modelo.

El siguiente modelo se basa en la metodología de descomposición Oaxaca-Blinder citada por los autores Blinder (1973) y Jann (2008), en donde se estiman dos regresiones lineales del salario en función a las características observables como: género, edad, nivel de escolaridad, experiencia, entre otras; una regresión para cada uno de los dos sectores.

$$\beta_{F1} - \beta_{inf1}$$

$$\beta_{F2}X - \beta_{inf2}Z = \beta_{F2}(X - Z) + Z(\beta_{F2} - \beta_{inf2})$$

Donde:

W_F = logaritmo natural del ingreso laboral formal

W_{inf} = logaritmo natural del ingreso laboral informal

X = características observables de los trabajadores formales

Z = características observables de los trabajadores informales

β_{inf1} = coeficiente, intercepción con el eje Y

β_{inf2} = pagos de dotaciones, mide la proporción de la variación del salario informal respecto a características observables de los trabajadores informales

β_{F1} = coeficiente, intercepción con el eje Y

β_{F2} = pagos de dotaciones, mide la proporción de la variación del salario formal respecto a características observables de los trabajadores formales

μ = errores

Se medirán las diferencias salariales a partir de:

$$R = \beta_{F1} + \beta_{F2}X - (\beta_{inf1} + \beta_{inf2}Z) = E + C + U$$

Con base en la ecuación anterior, las diferencias están explicadas en parte por la diferencia de las características observables entre el sector formal e informal, tales como: género, edad, nivel de escolaridad, experiencia, entre otras. Mientras que otra parte está explicada por el pago de estas observaciones (coeficientes). La última parte es explicada por la interacción entre las características observables y los coeficientes.

Es de esta manera que las primas diferenciales entre los salarios de los trabajadores del sector formal e informal están dadas por:

En donde E es la proporción diferencial que se le atribuye a las dotaciones, mientras que C es la porción de diferencias que se le atribuye a los coeficientes y U es la porción sin explicación del diferencial; entonces la suma de C y U es la porción de las diferencias atribuible a la discriminación.

3. Resultados y discusión

Modelo de Heckman

Se aplicó el modelo de Heckman a la submuestra para corroborar si realizando una estimación en dos etapas se logra una mayor explicación del ingreso salarial para el sector formal e informal, debido a que se consideran distintas variables que determinan, por un lado, probabilidad de que un individuo participe en el sector formal o en el sector informal, y por el otro, si esta probabilidad es estadísticamente significativa en un modelo de MCO. El programa econométrico utilizado para esta estimación es Gretl.

Tabla 1

Modelo de Heckit, estimación en dos etapas

Variable dependiente:	Ingreso Salarial			
Variable de selección:	Formal-Informal			
	Coeficiente	Desv. Típica	z	Valor p
Const	-3249.76	1107.03	-2.936	0.0033 ***
Nivel de escolaridad aprobado	1019.67	1107.03	7.179	7.01E-13 ***
Años de experiencia laboral	91.096	10.9168	8.345	7.15E-17 ***
Años de experiencia laboral ^2	-2.12232	0.35642	-5.955	2.61E-09 ***
Género	453.086	115.38	3.927	8.60E-05 ***
Edad de las personas (años cumplidos)	42.8846	4.7349	9.057	1.34E-19 ***
Etnicidad	469.981	137.584	3.416	0.0006 ***
Lambda	646.504	436.67	1.481	0.1387
Ecuación de selección				
Const	-3.70721	0.129148	-28.71	3.29E-181 ***
Género	0.346395	0.0313093	11.06	1.88E-28 ***
Etnicidad	0.372091	0.0344987	10.79	4.02E-27 ***
Alfabetismo	0.831781	0.0784661	10.6	4.96E-26 ***
Años de estudios aprobados	0.157098	0.00448392	35.04	6.41E-269 ***
Edad de las personas (años cumplidos)	0.0106503	0.00129574	8.219	2.04E-16 ***
Años de experiencia laboral	0.00262962	0.00430876	0.6103	0.5417
Años de experiencia laboral ^2	-0.000622588	0.00011073	-5.622	1.88E-08 ***
Sigma 1925.652	Rho	0.335733		

Fuente: estimación propia con base en INE (2018)

Según el método de Heckit, en los resultados, el valor de probabilidad asociado al «test t» con hipótesis nula $H_0: \text{Coef} = 0$, se indica con *** un $p < 0.02$; ** un $p < 0.05$; y * para un $p < 0.1$ (Rizzo, 2004).

Lo que interesa en esta estimación es la significancia estadística de la variable Lambda: como se planteó anteriormente, si (λ) resulta estadísticamente significativa, la muestra es considerada no aleatoria y las estimaciones estarían sesgadas de no incluir (λ) como variable explicativa en el modelo.

Los resultados del modelo de Heckman por estimación en dos etapas muestran que la variable de corrección (λ), razón inversa de Mills, es positiva y no es estadísticamente significativa al 0.05, 0.01 y 0.001, lo que significa que no se puede rechazar $H_0: \text{Coef} = 0$, por lo que no hay evidencia de un problema de selección muestral en los individuos que en las encuestas se autoseleccionaron como trabajadores en el sector formal y en el sector informal.

Varios de los valores definidos dentro de la ENEI 2018 fueron encontrados como *missing* o faltantes. Existen dos formas para solucionar este problema: sustitución de los datos faltantes por ceros, o la imputación de estos. En este caso, la sustitución por ceros no es la adecuada debido a que podría afectar en el cálculo de los resultados econométricos, por lo que fue necesario la imputación de estos. El procedimiento se realizó por medio del programa estadístico «R», identificando las observaciones faltantes y eliminando las mismas.

Este procedimiento se realizó en la variable ingresos de ambos sectores: para el modelo formal se encontraron 167 valores faltantes de 3398, lo que representa una imputación de 4.91 %, dejando la submuestra con 3230 observaciones. Para el modelo informal se encontraron 1019 valores faltantes, esto representa una imputación de 15.98 %, dejando la submuestra con 5343 observaciones. Estas imputaciones están por debajo del 20 % que recomienda la Comisión Económica para América Latina y el Caribe (Cepal) por medio de Medina y Galván (2007).

Modelo de Mincer

En el modelo de Mincer aplicado al sector formal, se tuvieron varias consideraciones teóricas propias del modelo econométrico mincenario, por lo que no se incluyeron otras variables parte de la teoría económica dentro de las regresiones. Estas variables serán tomadas en consideración en el modelo final de Oaxaca-Blinder, siempre y cuando coincidan con la teoría económica previamente analizada.

A continuación, se muestra una tabla de resumen de los resultados obtenidos de la regresión para el sector formal. El programa econométrico utilizado para la estimación es «R».

Tabla 2
Resumen del modelo de Mincer aplicado al sector formal

Call:								
lm(formula	=	log(P04C10)	~	P03A05AB+	P04C30+	l(P04C30^2),	data=	bd.f4.na2)
Residuals:								
	Min	1Q	Median	3Q	Max			
	-3.7259	-0.2178	0.0302	0.2724	1.8945			
Coefficients:								
	Estimate	Std. Error	t value	Pr(> t)				
(Intercept)	6.99500	0.02869	243.785	0.00000	***			
P03A05AB	0.05853	0.00187	31.383	0.00000	***			
P04C30	0.03652	0.00255	14.33	0.00000	***			
l(P04C30^2)	-0.00067	0.00008	-8.761	0.00000	***			
Signif.	codes:	0 '***'	0.001 '**'	0.01 '*'	0.05 '.'	0.1 ''	1	
Residual standard error:	0.4739	on 3226	degrees of freedom					
Multiple R-squared:	0.2932,	Adjusted R-squared:	0.2925					
F-statistic:	446	on 3	and 3226	DF,	p-value:	0.00000000		

Fuente: estimación propia con base en INE (2018)

En donde:

P04C10: ingreso salarial de los trabajadores formales

(Intercept): coeficiente, intercepción con el eje Y

P03A05AB: número de años de educación formal completada para los trabajadores formales

P04C30: años de experiencia laboral para los trabajadores formales

l(P04C30^2): años de experiencia laboral al cuadrado para los trabajadores formales

a. Nivel de escolaridad del sector formal

En lo referente al nivel de escolaridad de los individuos del sector formal, el valor del coeficiente mostró un resultado de 0.058. Considerando el sentido logarítmico del modelo, este valor indica que una persona que forma parte de la población económicamente activa empleada en el sector formal, obtiene un ingreso 5.8 % mayor que el de una persona dentro de la población económicamente activa y en el mismo sector, con un nivel educativo inferior.

b. Experiencia laboral del sector formal

Con respecto a la experiencia laboral, el valor del coeficiente es de 0.036, lo que significa que una persona de la población económicamente activa en el sector formal tiene un ingreso 3.6 % mayor que el de una persona con las mismas características, pero con un año menos de experiencia laboral.

En el caso de la experiencia laboral al cuadrado, el valor del coeficiente es de -0.00067: esto indica que a medida que la experiencia aumenta, su relación con el ingreso se torna decreciente. De acuerdo con lo anterior, si bien la experiencia laboral tiene una relación directamente proporcional con el ingreso laboral de los individuos de la población económicamente activa, de igual manera comienza a decaer a medida que transcurren los años laborales, y esta caída es de un aproximado de 0.067 % por cada año de experiencia laboral adicional.

A continuación, se muestra una tabla de resumen de los resultados obtenidos del modelo de Mincer para el sector informal. El programa econométrico utilizado para la estimación es «R».

Tabla 3
Resumen del modelo de Mincer aplicado al sector informal

Call:								
lm(formula	=	log(P04C10B)	~	P03A05AB+	P04C30+	I(P04C30^2),	data=	bd.f4.na2)
Residuals:								
	Min	1Q	Median	3Q	Max			
	-5.1829	-0.488	0.0602	0.6004	4.0347			
Coefficients:								
	Estimate	Std. Error	t value	Pr(> t)				
(Intercept)	6.44200	0.03627	177.64	0.00000	***			
P03A05AB	0.04955	0.00279	17.78	0.00000	***			
P04C30	0.02765	0.00273	10.13	0.00000	***			
I(P04C30^2)	-0.00067	0.00006	-11.73	0.00000	***			
--- Signif.	codes:	0 '***'	0.001 '**'	0.01 '*'	0.05 '.'	0.1 ''	1	
Residual	standard error:	0.9334	on	5339	degrees	of	freedom	
Multiple	R-squared:	0.08457,	Adjusted	R-squared:	0.08405			
F-statistic:	164.4	on	3	and	5339	DF,	p-value:	0.00000000

Fuente: estimación propia con base en INE (2018)

En donde:

P04C10B: ingreso salarial de los trabajadores informales

(Intercept): coeficiente, intercepción con el eje Y

P03A05AB: número de años de educación formal completada para los trabajadores informales

P04C30: años de experiencia laboral para los trabajadores informales

I(P04C30^2): años de experiencia laboral al cuadrado para los trabajadores informales

c. Nivel de escolaridad del sector informal

En consideración al nivel de escolaridad de los individuos del sector informal, el valor del coeficiente mostró un resultado de 0.0495. Teniendo en cuenta el sentido logarítmico del modelo, este valor indica que una persona que forma parte de la población económicamente activa empleada en el sector informal obtiene un ingreso 4.95 % mayor que el de una persona dentro de la población económicamente activa y en el mismo sector, con un nivel de escolaridad inferior.

d. Experiencia laboral del sector informal

Con respecto a la experiencia laboral, el valor del coeficiente es de 0.027, lo que significa que una persona de la población económicamente activa en el sector informal tiene un ingreso 3.6 % mayor que el de una persona con las mismas características, pero con un año menos de experiencia laboral.

En el caso de la experiencia laboral al cuadrado, el valor del coeficiente es de -0.00067, relativamente el mismo coeficiente del sector formal: esto indica que a medida que la experiencia aumenta, su relación con el ingreso se torna decreciente. De acuerdo con lo anterior, si bien la experiencia laboral tiene una relación directamente proporcional con el ingreso laboral de los individuos de la población económicamente activa, de igual manera comienza a decaer a medida que transcurren los años laborales, y esta caída es de un aproximado de 0.067 % por cada año de experiencia laboral adicional.

e. Modelo de Oaxaca-Blinder

A continuación, se muestra una tabla de resumen de los resultados obtenidos de la regresión para el sector formal. El programa econométrico utilizado para la estimación es «R».

Tabla 4
Resumen del modelo de Oaxaca-Blinder aplicado al sector formal

Call:								
ln(formula	=	log(P04C10)	~	P03A05A+	PPA03+	PPA02+	P04C30+	l(P04C30^2)+
		ETNICIDAD	data=	bd.f4.na2)				
Residuals:								
	Min	1Q	Median	3Q	Max			
	-3.6645	-0.2179	0.0308	0.2688	1.8233			
Coefficients:								
	Estimate	Std. Error	t value	Pr(> t)				
(Intercept)	6.70392	0.05583	120.088	0.00000	***			
P03A05APreprimaria	0.18806	0.26610	0.707	0.48000				
P03A05APrimaria	0.30431	0.04718	6.449	0.00000	***			
P03A05ABasicos	0.48108	0.04909	9.799	0.00000	***			
P03A05ADiversificado	0.70047	0.04661	15.027	0.00000	***			
P03A05ASuperior	1.01077	0.04856	20.813	0.00000	***			
P03A05AMaestria	1.51409	0.10118	14.964	0.00000	***			
P03A05ADoctorado	1.36938	0.32429	4.223	0.00002	***			
PPA03	0.00901	0.00086	10.419	0.00000	***			
PPA02hombre	0.11812	0.01754	6.734	0.00000	***			
P04C30	0.02581	0.00263	9.801	0.00000	***			
l(P04C30^2)	-0.00060	0.00007	-8.184	0.00000	***			
ETNICIDADladino	0.18871	0.02163	8.723	0.00000	***			

Signif.	codes:	0 '***'	0.001 '**'	0.01 '*'	0.05 '.'	0.1 ''	1	
Residual	standard	error:	0.4539	on	3217	degrees	of	freedom
Multiple	R-squared:	0.3535,	Adjusted	R-squared:	0.3511			
F-statistic:	146.6	on	12	and	3217	DF,	p-value:	0.00000000

Fuente: estimación propia con base en INE (2018)

En donde:

P04C10: ingreso salarial de los trabajadores formales

(Intercept): coeficiente, intercepción con el eje Y

P03A05APreprimaria: variable dicotómica, toma el valor de 1 si el individuo aprobó el nivel de preprimaria y 0 si el individuo no aprobó ningún nivel académico

P03A05APrimaria: variable dicotómica, toma el valor de 1 si el individuo aprobó el nivel primario y 0 si el individuo no aprobó ningún nivel académico

P03A05ABásicos: variable dicotómica, toma el valor de 1 si el individuo aprobó el nivel básico y 0 si el individuo no aprobó ningún nivel académico

P03A05ADiversificado: variable dicotómica, toma el valor de 1 si el individuo aprobó el nivel de diversificado y 0 si el individuo no aprobó ningún nivel académico

P03A05ASuperior: variable dicotómica, toma el valor de 1 si el individuo aprobó el nivel superior y 0 si el individuo no aprobó ningún nivel académico

P03A05AMaestría: variable dicotómica, toma el valor de 1 si el individuo aprobó el nivel de maestría y 0 si el individuo no aprobó ningún nivel académico

P03A05ADoctorado: variable dicotómica, toma el valor de 1 si el individuo aprobó el nivel de doctorado y 0 si el individuo no aprobó ningún nivel académico

P04C30: años de experiencia laboral para los trabajadores formales

I(P04C30²): años de experiencia laboral al cuadrado para los trabajadores formales

PPA03: edad de las personas (años cumplidos)

PPA02Hombre: variable dicotómica, toma el valor de 1 si es hombre y 0 si es mujer

ETNICIDADLadino: variable dicotómica, toma el valor de 1 si es ladino y 0 si es indígena

a. Nivel de escolaridad del sector formal

Para el nivel primario, el valor del coeficiente mostró un resultado de 0.3043: este valor indica que una persona que forma parte de la población económicamente activa que aprobó el nivel primario obtiene un ingreso 30.43 % mayor que el de una persona con las mismas características, pero que no aprobó ningún nivel académico.

En el nivel básico, el valor del coeficiente mostró un resultado aproximado de 0.4811: este valor indica que una persona que forma parte de la población económicamente activa que aprobó el nivel básico percibe un ingreso 48.11 % superior al de una persona con las mismas características, pero que no aprobó ningún nivel académico.

Con respecto al nivel de diversificado, el valor del coeficiente mostró un resultado aproximado de 0.7005: este valor indica que una persona que forma parte de la población económicamente activa que aprobó el nivel de diversificado cobra un ingreso 70.05 % superior al de una persona con las mismas características, pero que no aprobó ningún nivel académico.

En relación con el nivel superior, el valor del coeficiente mostró un resultado aproximado de 1.0108: este valor indica que una persona que forma parte de la población económicamente activa y que aprobó el nivel superior recibe un ingreso 101.08 % mayor al de una persona con las mismas características, pero que no aprobó ningún nivel académico.

En el caso del nivel de maestría, el valor del coeficiente mostró un resultado aproximado de 1.5141: este valor indica que una persona que forma parte de la población económicamente activa y que aprobó el nivel de maestría obtiene un ingreso 151.41 % mayor que el de una persona con las mismas características, pero que no aprobó ningún nivel académico.

Para el nivel de doctorado, el valor del coeficiente mostró un resultado aproximado de 1.3694: este valor indica que una persona que forma parte de la población económicamente activa y que aprobó el nivel de doctorado percibe un ingreso 136.94 % superior al de una persona con las mismas características, pero que no aprobó ningún nivel académico.

b. Experiencia laboral del sector formal

Con respecto a la experiencia laboral, el valor del coeficiente es de 0.0258, lo que significa que una persona de la población económicamente activa en el sector formal tiene un ingreso 2.58 % mayor al de una persona con las mismas características, pero con un año menos de experiencia laboral.

En el caso de la experiencia laboral al cuadrado, el valor del coeficiente es de -0.00060: esto indica que, a medida que la experiencia aumenta, su relación con el ingreso se torna decreciente. De acuerdo con lo anterior, si bien la experiencia laboral tiene una relación directamente proporcional con el ingreso laboral de los individuos de la población económicamente activa, de igual manera comienza a decaer a medida que transcurren los años laborales, y esta caída es de un aproximado de 0.06 % por cada año de experiencia laboral adicional.

c. Variable con respecto al género del sector formal

Los individuos de la población económicamente activa empleados en el sector formal que son hombres perciben un ingreso 11.81 % superior al de las mujeres.

d. Variable con respecto a la etnicidad del sector formal

Los individuos de la población económicamente activa empleados en el sector formal que se autodenominaron ladinos ganan un 18.87 % más que las personas que se autodenominaron indígenas.

e. Variable edad del sector formal

Los individuos parte de la población económicamente activa dentro del sector formal de la economía aumenta su ingreso por cada año en un 0.90 %.

A continuación, se muestra una tabla de resumen de los resultados obtenidos de la regresión para el sector informal. El programa econométrico utilizado para la estimación es «R».

Tabla 5
Resumen del modelo de Oaxaca-Blinder aplicado al sector informal

Call:								
ln(formula	=	log(P04C10B)	~	P03A05A+	PPA03+	PPA02+	P04C30+	((P04C30^2)+
		ETNICIDAD	data=	bd.f4.na2)				
Residuals:								
	Min	1Q	Median	3Q	Max			
	-5.219	-0.451	0.069	0.5524	4.0983			
Coefficients:								
	Estimate	Std. Error	t value	Pr(> t)				
(Intercept)	6.06300	0.05456	111.122	0.00000	***			
P03A05APreprimaria	0.31310	0.19190	1.632	0.10300				
P03A05APrimaria	0.21130	0.03574	5.912	0.00000	***			
P03A05ABasicos	0.42910	0.04653	9.221	0.00000	***			
P03A05ADiversificado	0.76180	0.04676	16.291	0.00000	***			
P03A05ASuperior	0.96740	0.11330	8.538	0.00000	***			
PPA03	0.00747	0.00107	7.012	0.00000	***			
P04C30	0.01973	0.00288	6.848	0.00000	***			
l(P04C30^2)	-0.00066	0.00006	-11.778	0.00000	***			
PPA02hombre	0.25590	0.02700	9.479	0.00000	***			
ETNICIDADladino	0.24850	0.02718	9.141	0.00000	***			

Signif.	codes:	0 '***'	0.001 '**'	0.01 '*'	0.05 '.'	0.1 ''	1	
Residual	standard	error:	0.9085	on	5332	degrees	of	freedom
Multiple	R-squared:	0.1337,	Adjusted	R-squared:	0.1321			
F-statistic:	82.32	on	10	and	5332	DF,	p-value:	0.00000000

Fuente: estimación propia con base en INE (2018)

En donde:

P04C10B: ingreso salarial de los trabajadores informales

(Intercept): coeficiente, intercepción con el eje Y

P03A05APreprimaria: variable dicotómica, toma el valor de 1 si el individuo aprobó el nivel de preprimaria y 0 si el individuo no aprobó ningún nivel académico

P03A05APrimaria: variable dicotómica, toma el valor de 1 si el individuo aprobó el nivel primario y 0 si el individuo no aprobó ningún nivel académico

P03A05ABásicos: variable dicotómica, toma el valor de 1 si el individuo aprobó el nivel básico y 0 si el individuo no aprobó ningún nivel académico

P03A05ADiversificado: variable dicotómica, toma el valor de 1 si el individuo aprobó el nivel de diversificado y 0 si el individuo no aprobó ningún nivel académico

P03A05ASuperior: variable dicotómica, toma el valor de 1 si el individuo aprobó el nivel superior y 0 si el individuo no aprobó ningún nivel académico

P04C30: años de experiencia laboral para los trabajadores informales

l(P04C30^2): años de experiencia laboral al cuadrado para los trabajadores informales

PPA03: edad de las personas (años cumplidos)

PPA02Hombre: variable dicotómica, toma el valor de 1 si es hombre y 0 si es mujer

ETNICIDADLadino: variable dicotómica, toma el valor de 1 si es ladino y 0 si es indígena

a. Nivel de escolaridad del sector informal

Para el nivel primario, el valor del coeficiente mostró un resultado de 0.2113: este valor indica que una persona que forma parte de la población económicamente activa que aprobó el nivel primario obtiene un ingreso 21.13 % mayor al de una persona con las mismas características, pero que no aprobó ningún nivel académico.

En el nivel básico, el valor del coeficiente mostró un resultado aproximado de 0.4291: este valor indica que una persona que forma parte de la población económicamente activa empleada en el sector informal y que aprobó el nivel básico, percibe un ingreso 42.91 % superior que el de una persona con las mismas características, pero que no aprobó ningún nivel académico.

Con respecto al nivel de diversificado, el valor del coeficiente mostró un resultado aproximado de 0.7618: este valor indica que una persona que forma parte de la población económicamente activa empleada en el sector informal y que aprobó el nivel de diversificado, cobra un ingreso 76.18 % superior que el de una persona con las mismas características, pero que no aprobó ningún nivel académico.

En relación con el nivel superior, el valor del coeficiente mostró un resultado aproximado de 0.9674: este valor indica que una persona que forma parte de la población económicamente activa empleada en el sector informal y que aprobó el nivel superior, recibe un ingreso 96.74 % mayor que el de una persona con las mismas características, pero que no aprobó ningún nivel académico.

b. Experiencia laboral del sector informal

Con respecto a la experiencia laboral, el valor del coeficiente es de 0.0197, lo que significa que una persona de la población económicamente activa en el sector informal tiene un ingreso 1.97 % mayor que el de una persona con las mismas características, pero con un año menos de experiencia laboral.

En el caso de la experiencia laboral al cuadrado, el valor del coeficiente es de -0.0006: esto indica que, a medida que la experiencia aumenta, su relación con el ingreso se torna decreciente. De acuerdo con lo anterior, si bien la experiencia laboral tiene una relación directamente proporcional con el ingreso laboral de los individuos de la población económicamente activa, de igual manera comienza a decaer a medida que transcurren los años laborales, y esta caída es de un aproximado de 0.06 % por cada año de experiencia laboral adicional.

c. Variable con respecto al género del sector informal

Los individuos de la población económicamente activa empleados en el sector informal que son hombres perciben un ingreso 25.59% superior al de las mujeres.

d. Variable con respecto a la etnicidad del sector informal

Los individuos de la población económicamente activa empleados en el sector informal que se autodenominaron ladinos ganan un 24.85 % más que las personas que se autodenominaron indígenas.

e. Variable edad del sector informal

Los individuos parte de la población económicamente activa dentro del sector informal de la economía indican un aumento en su ingreso por cada año de edad, en un 0.74 %.

Según la descomposición salarial Oaxaca-Blinder, se estima la brecha salarial entre los trabajadores del sector formal e informal de la población económicamente activa del país. Esta brecha puede deberse a las diferencias grupales, específicamente en los determinantes salariales mencionados en las regresiones anteriores (Marek, 2018).

Tabla 6

Resumen de la descomposición salarial Oaxaca-Blinder

>	Resultados
Grupo.A	
[1]	3646.01
Grupo.B	
[1]	1724.68
Diferencia	
[1]	-0.9989385

Fuente: estimación propia con base en INE (2018)

El componente Oaxaca resultante indica que el salario medio de los trabajadores formales (grupo A) es de Q3646.01, y para los trabajadores informales es de Q1724.68, dejando así la brecha salarial en un 99.89 %, explicada por la descomposición Blinder-Oaxaca (Marek, 2018).

Tabla 7
Resumen de la descomposición triple Oaxaca-Blinder

Resultados	
coeficientes(dotaciones)	se(dotaciones)
0.32628173	0.01556128
Coeficientes(coeficientes)	se(coeficientes)
0.72247158	0.02292026
Coeficientes(interacción)	se(interacción)
-0.04981479	0.02256621

Fuente: estimación propia con base en INE (2018)

Se le llama descomposición triple debido a que muestra el análisis en donde E es la proporción diferencial que se le atribuye a las dotaciones, mientras que C es la porción de diferencias que se le atribuye a los coeficientes y U es la porción sin explicación del diferencial.

Los resultados de la triple descomposición indican que, de la diferencia del 99.89 %, se le atribuye aproximadamente el 32.63 % de la brecha a las diferencias en las dotaciones (género, edad, experiencia, escolaridad y etnicidad); mientras el 72.25 % de la brecha se atribuye a las diferencias en los coeficientes, es decir, por los pagos de las dotaciones. Esto significa que mide el cambio en los salarios de los trabajadores informales cuando se le aplican los coeficientes de los trabajadores formales a las características de los informales: por ejemplo, la remuneración que reciben los trabajadores formales por un año más de escolaridad es superior al pago que recibiría un trabajador informal (Manco Rivera, 2016). El 0.05 % restante se le atribuye a la interacción de los coeficientes y las dotaciones. Para comprender la discriminación, se adjunta el siguiente resumen de la descomposición.

Tabla 8
Resumen de las primas diferenciales

Primas diferenciales		
E =	Proporción atribuible a las dotaciones	32.63 %
C =	Proporción atribuible a los coeficientes	72.25 %
U =	Proporción sin explicación del diferencial	-0.05 %
C+U =	Proporción atribuible a la discriminación	72.20 %

Fuente: estimación propia con base en INE (2018)

La proporción atribuible a la discriminación es de un 72.20 % a favor de los trabajadores del sector formal: esto quiere decir *ceteris paribus* que un individuo, solo por trabajar en el sector informal, percibirá un ingreso 72.20 % menor al que percibiría si trabajara en el sector formal.

La razón principal de esta brecha se debe a que los trabajadores formales poseen mejores características: por ejemplo, en el nivel de escolaridad de la población económicamente activa en el sector informal, el 16.52 % del total no cursó ningún nivel académico, mientras que en el sector formal solo el 3.18% no cursó ningún nivel académico. En el caso de la experiencia laboral, en el sector formal los trabajadores reportan un 62.48 % más de experiencia que los trabajadores informales, lo que genera una mejor remuneración. En las regresiones, se refleja que un año más de educación o un año más de experiencia laboral tiene mayor incidencia en el ingreso laboral para el sector formal que para el sector informal.

Tabla 9
P-value

Simple t-test		
datos:	Resultados descomposición Oaxaca-Blinder	
t= 2.834,	df=53	p-value=0.006491
Hipótesis alternativa	La verdadera medida no es igual a 0	
95 por ciento	confianza	intervalo:
0.04961002	0.28989544	
Estimaciones simples:		
Media de x	0.1697527	

Fuente: estimación propia con base en INE (2018)

Se estimó el *p-value* para la descomposición Oaxaca-Blinder, tomando como base la regresión del grupo A como el sector formal, y la regresión del grupo B, que es el sector informal.

El *p-value* es de 0.006491, menor al 5 % en las variables observadas, por lo que la hipótesis nula se rechaza, lo que significa que los salarios medios y la brecha salarial son estadísticamente significativos.

4. Conclusiones

Se acepta la hipótesis de investigación, lo que significa que el nivel de ingreso de los trabajadores del sector formal de la economía en Guatemala es significativamente mayor con respecto al nivel de ingreso de los trabajadores en el sector informal, con un nivel de confianza del 95 %.

Las diferencias principales en el nivel de ingreso de los trabajadores del sector formal e informal son: el nivel de escolaridad, género, etnicidad y la experiencia laboral. En el sector formal, se muestra un nivel superior de escolaridad que el del sector informal, lo que genera una mayor remuneración, además que un año de educación adicional en la mayoría de los niveles en el sector formal tiene mayor incidencia en el ingreso que un año adicional en el nivel de educación en el sector informal.

Con respecto a la variable género, ser hombre incide más en el ingreso de los trabajadores del sector informal que en el sector formal; lo mismo sucede con la variable etnicidad: en el sector informal, ser ladino tiene una mayor incidencia en el ingreso que en el sector formal. Caso contrario con la experiencia laboral: un año de experiencia adicional tiene mayor incidencia en el ingreso para el sector formal que para el sector informal.

En el sector formal, las características que tienen mayor incidencia en el ingreso son: el nivel de escolaridad, la variable género y la variable etnicidad. En el nivel de escolaridad, haber aprobado el nivel primario representa un ingreso 30.43 % mayor que el no haberlo aprobado; 17.68 % más el haber aprobado el nivel básico; en el nivel de diversificado, aumenta un 21.94 % el ingreso; tener un grado superior aumenta el ingreso en un 31.03 %; en el nivel de maestría, aumenta el ingreso un 50.33 %; y en el caso de doctorado, sucede lo contrario al patrón visto en los demás niveles educativos, pues en lugar de aumentar el ingreso, tener dicho grado disminuye el ingreso en comparación al nivel de maestría en un 14.47 %.

Para la variable género, ser hombre genera un ingreso 11.81 % mayor que ser mujer, y para la variable etnicidad, ser ladino incide en un 18.87 % más en el ingreso que ser indígena.

En el sector informal, las características que tienen mayor incidencia en el ingreso son: el nivel de escolaridad, la variable género y la variable etnicidad. Haber aprobado el nivel primario representa un ingreso 21.13 % mayor que no haberlo aprobado; para el nivel básico aumenta un 21.78 % el ingreso; 33.27 % más en el ingreso por haber aprobado el nivel de diversificado; y para el nivel superior, aumenta el ingreso en un 20.56 %.

Para la variable género, ser hombre representa un ingreso 25.59 % mayor que el ingreso del género femenino; y para la variable etnicidad, ser ladino genera un ingreso 24.85 % mayor que ser indígena.

El ingreso promedio de los trabajadores del sector formal obtenido de la descomposición Oaxaca-Blinder es de Q3646.01. El ingreso promedio de los trabajadores del sector informal obtenido de la descomposición Oaxaca-Blinder es de Q1724.68.

La brecha salarial explicada por la descomposición Oaxaca-Blinder es de un 99.89 %. El 32.63 % es explicado por las diferencias en las dotaciones (género, edad, experiencia laboral, escolaridad y etnicidad), el 72.25 % se le atribuye a las diferencias en los coeficientes, y el 0.05 % restante se explica por la interacción entre los coeficientes y las dotaciones.

El componente de discriminación por la descomposición Oaxaca-Blinder es de un 72.20 %, lo cual significa que un individuo, únicamente por trabajar en el sector formal, percibirá un ingreso 72.20 % mayor.

Referencias

- Blinder, A. (1973). *Wage Discrimination Reduced Form and Structural Estimates*. *Journal of Human Resources*. Wisconsin: University of Wisconsin Press.
- Heckman, J. (1979). Sample Selection Bias as a Specification Error. *Econometrica Journal*, 47(1), 153-161. doi: 10.2307/1912352
- Instituto Nacional de Estadística (INE). (2018). *Encuesta Nacional de Empleo e Ingresos (ENEI)*. https://www.ine.gob.gt/sistema/uploads/2019/07/05/publicacion_ENEI_2_2018.pdf
- Jann, B. (2008). The Blinder–Oaxaca decomposition for linear regression models. *The Stata Journal*, 8(4), 453-479.
- Manco Rivera, C. (2016). *Determinantes de la brecha salarial entre el empleo formal e informal en Colombia*. Medellín, Colombia: Escuela de Economía y Finanzas, Universidad EAFIT. <https://core.ac.uk/download/pdf/81651364.pdf>
- Marek, H. (2018). *Oaxaca: Blinder-Oaxaca Decomposition in R. R package version*. <https://CRAN.R-project.org/package=oaxaca>
- Medina, F. y Galván, M. (2007). *Imputación de datos: teoría y práctica*. Santiago de Chile: Comisión Económica para América Latina y el Caribe (Cepal).
- Metzger, R. S. (2006). *Rendimiento a la escolaridad en México: un enfoque de efecto tratamiento*. Puebla, México: Banco de México.
- Mincer, J. (1974). *Experience and Earnings*. New York: National Bureau of Economic Research.
- Perodi, S., Ramírez, I. y Thompson, J. (2017). *Tasas de retorno de la inversión en República Dominicana (2000-2005)*. Washington: Banco Interamericano de Desarrollo.
- Rizzo, J. I. (2016). *Un análisis exploratorio de los determinantes del gasto turístico en recreación y alimentación en Mar del Plata* (Tesis doctoral). Universidad Nacional de Mar del Plata, Argentina.